A Proteome-Wide Network Analysis of the Interrelationships between Kinases, Phosphatases, and their Putative Substrates in Saccharomyces cerevisiae

Dissertation zur Erlangung der
naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der
Mathematisch-naturwissenschaftlichen Fakultät
der
Universität Zürich

von
Stefanie Wanka
aus Deutschland

Promotionskomitee:
Prof. Dr. Christian von Mering
(Vorsitz und Leiter der Dissertation)
Dr. Anne-Claude Gavin
Dr. Matthias Gstaiger
Prof. Dr. Andreas Wagner

Zürich 2011

ABSTRACT

Living systems are formed and defined by the complex interplay of a variety of biological networks, implicating the need of an accurate but flexible regulation mechanism. Phosphorylation networks comprising the interrelationships between kinases, phosphatases, and their substrates take on a leading part in the cellular regulation system, and are of fundamental importance for the function and robustness of essentially all biological activities including cell division, signal transduction, metabolism, and cell motility.

Despite the undoubted relevance of phosphorylation networks, rather little is known about their architecture and organization so far. In fact, due to technological developments, the amount of phosphoproteome data has been growing in recent years, resulting in the listing and characterization of phosphoproteins and their sites of phosphorylation in a variety of organisms. Nevertheless, the dependencies between kinases, phosphatases and phosphoproteins are mostly not understood and only for few kinases and phosphatases their immediate downstream targets are known. Linking the kinome and phosphatome to the phosphoproteome still remains a challenging task.

In this study I analyzed and explored the raw data from a collaborative project which used label–free, quantitative phosphoproteomics in order to determine the relationships between 97 kinases, 27 phosphatases, and more than 1000 phosphoproteins in the yeast *Saccharomyces cerevisiae*. We systematically detected and identified phosphopeptides that showed a significant and reproducible change in their abundance upon the removal (or inactivation) of each of the kinases or phosphatases, thereby describing the first system–wide *in vivo* phosphorylation network of *Saccharomyces cerevisiae*. Our analysis revealed a high degree of interconnectedness within the network, as the inactivation of most kinases and phosphatases did not only affect potential direct targets but large parts of the phosphorylation network. This outcome argues for a considerable robustness in the architecture of the phosphorylation–dependent control machinery and supports the idea that signaling pathways do not act in a simple linear manner, an observation of consequence for drug development and for describing organismal homeostasis.

ZUSAMMENFASSUNG

Lebende Systeme beruhen auf dem komplexen Zusammenspiel verschiedener zellulärer Netzwerke, ein Mechanismus, der sowohl präzise als auch flexible Regulierungsmöglichkeiten erfordert. Phosphorylierungsnetzwerke beschreiben die Zusammenhänge zwischen Kinasen, Phosphatasen und ihren Substraten. Sie spielen eine wichtige Rolle im zellulären Kontrollsystem und sind für den Ablauf im Grunde aller biologischen Prozesse, einschliesslich Zellteilung, Signalübertragung, Metabolismus und Zellmotilität, von grundlegender Bedeutung.

Obwohl die Bedeutsamkeit von Phosphorylierungsnetzwerken unbestritten ist, ist über deren Architektur und Gestalt bis zum heutigen Zeitpunkt nur wenig bekannt. So werden zwar aufgrund technischer Fortschritte immer grössere Mengen an Phosphoproteom–Daten erzeugt, was zu einem stetigen Anstieg bekannter Phosphoproteine und ihren Phosphorylierungsstellen führt, allerdings sind die Zusammenhänge und Abhängigkeiten der Kinasen, Phosphatasen und Substrate untereinander zum grossen Teil immer noch unbekannt. Bis dato konnte nur für eine kleine Anzahl von Kinasen die direkten Zielproteine identifiziert werden und die Verknüpfung von Kinom und Phosphatom mit dem Phosphoproteom stellt immer noch eine grosse Herausforderung dar.

In der vorliegenden Arbeit analysierte und untersuchte ich die mittels quantitativer Phosphoproteomik erzeugten Rohdaten eines gemeinschaftlichen Projekts mit dem Ziel, die Beziehungen zwischen 97 Kinasen, 27 Phosphatasen und über 1000 Phosphoproteinen in der Hefe *Saccharomyces cerevisiae* zu beschreiben und zu erläutern. Systematisch identifizierten wir Phosphopeptide, die, als Reaktion auf die Inaktivierung einer bestimmten Kinase oder Phosphatase, eine signifikante und reproduzierbare Änderung in ihrer Abundanz aufwiesen. Auf diese Weise konnten wir das erste systemweite *in vivo* Phosphorylierungsnetzwerk von *Saccharomyces cerevisiae* beschreiben. Unsere Analyse machte dabei den hohen Grad an Vernetzung und Interaktion innerhalb des Netzwerks deutlich, da die Inaktivierung der meisten Kinasen und Phosphatasen nicht nur nachgeschaltete Phosphoproteine, sondern oft grössere Teile des Netzwerks beeinflusste. Dieses Ergebnis weist auf eine erhebliche Robustheit im phosphorylierungsbasierten Regulierungssystem der Zelle hin und bestärkt die Annahme, dass Signalübertragungswege keinesfalls in linearer Art und Weise agieren. Eine Erkenntnis, die auch hinsichtlich der Entwicklung von Medikamenten und der Beschreibung der Homeostasis von Organismen von wichtiger Bedeutung ist.

ACKNOWLEDGMENTS

First and foremost I would like to thank my advisor Christian von Mering. Christian *lives* science. I am deeply grateful for his enthusiastic supervision and continuous support throughout my PhD. Christian always has an open door to questions and problems, both literally and figuratively, a rare but very desirable quality for a supervisor. Christian, thanks a lot for your patient help and excellent guidance!

Next I would like to thank my Thesis Committee, Dr. Anne-Claude Gavin, Dr. Matthias Gstaiger, and Prof. Andreas Wagner for their advice and support throughout the years and for all the stimulating discussions.

Gratefulness beyond words to Bernd Bodenmiller. He is not only an excellent scientist but also a great person. It was a true pleasure to work with him.

Thanks a lot to Juliane Schulz, too. It is great to collaborate with her and she is always supportive and patient with a (former) non-metabolomics person like me.

I would also like to thank Gerlinde Reim, with whom I collaborate on a second project and who has broadened my mind on developmental biology.

Many thanks to all present and past von Mering group members! We really have a great team atmosphere with fun and laughs which makes working life much easier! Special thanks to the 'old core': Manuel Stark, with whom I could communicate without words, Samuel Chaffron and his irresistible French-Belgian charm, Manuel Weiss and his enviable calmness. Guys, I am missing you a lot! Big hugs also to Milan Simonovic and Alexander Roth! And special thanks to Gabi, who was a great help in creating these very nice phosphosite–metabolite correlation plots.

1000 thanks to my friends for emotional support and motivation, for distraction and fun and enjoying a life beyond science. Especially to Eva and Regina, who I have known almost all my life, to Andy, who has supported me in making important decisions and has been through hard times with me, to Marko, the one and only, guide, friend and realist with having the rare gift of addressing also unpleasant truths in the right way, and to Erica, who is like a sister to me.

Last but not least I would like to thank my family for their love, support, encouragement and always believing in me. I am very grateful for their understanding and for putting up with my moodiness in hard and difficult times!

CONTENTS

ACRONYMS

ATP	adenosine triphosphate
AUC	area under the curve
DNA	deoxyribonucleic acid
ESI	electrospray ionization
FIE–MS	(direct) flow injection electrospray mass spectrometry
GC–MS	gas chromatography mass spectrometry
GO	gene ontology
HTP	high–throughput
ICAT	isotope–coded affinity tag
IMAC	immobilized metal affinity chromatography
LC–MS	liquid chromatography mass spectrometry
LC–MS/MS	liquid chromatography coupled to tandem mass spectrometry
LTP	low–throughput
MS	mass spectrometry
MS1	precursor ion scan of the mass spectrometer
MS2	MS/MS peptide fragment mass scan
MS/MS	tandom mass spectrometry
m/z	mass to charge [Da]
NMR	nuclear magnetic resonance
PPI	protein protein interaction
PTM	post–translational modification
RNAi	RNA interference
Rt	retention time
SILAC	stable isotope labeling by amino acids in cell culture
S/N	signal–to–noise

TF	transcription factor
TiO_2	titanium dioxide

Part I

INTRODUCTION AND CONTEXT

1 INTRODUCTION

Systems Biology is a broad field, comprising a wide range of biological research areas and incorporating both experimental and computational technologies [1, 2, 3, 4, 5]. While the classification of Systems Biology as an independent field of science has been somewhat controversial, it is clear that the term denotes a set of newly developed experimental, mathematical and computational methods applied to solving biological problems [6]. This points out that the rise of Systems Biology would not have been possible without technological advancements. Rather, it actually coincides with and is based on the invention and development of the so called 'omics' disciplines including genomics, transcriptomics, proteomics, and metabolomics.

The conceptual novelty of Systems Biology is also reflected by its lack of a concise definition. One of many working definitions is "the search for the syntax of biological information, that is, the study of the dynamic networks of interacting biological elements" [7]. Various 'ome' units (genome, transcriptome, proteome, etc.) constitute such biological elements. The concerted interplay of their underlying biological networks forms and defines all living systems. As a result, a complex framework arises that is subject to precise regulation, carried out by adequate regulatory networks.

In 1992, Edmond H. Fischer and Edwin G. Krebs were awarded with the Nobel Prize in Physiology or Medicine for their "discoveries concerning reversible protein phosphorylation as a biological regulatory mechanism" [8, 9]. Initially only characterized in carbohydrate metabolism [10], it soon became clear that phosphorylation is indeed one of the most prevalent mechanism by which the cell turns a reaction on or off [11]. Thus, together with the transcriptional network, the phosphorylation network is considered as one of the most prominent regulatory networks: Protein kinases and protein phosphatases mediate signals by phosphorylating and dephosphorylating their target (phospho–) proteins, respectively, hereby controlling a wide range of essential cellular processes including cell division, transcription, growth, and metabolic pathways [12, 13].

Although many of the individual players in the phosphorylation networks of various organisms are known, the topological landscapes of these networks are not. The signaling context for most phosphorylation sites continues to be unidentified, i.e. there is a big gap in linking kinases and phosphatases to their respective substrates [14]. For instance, information about the upstream kinases and phosphatases is available for only 13% of the known phosphorylation sites of the yeast *Saccharomyces cerevisiae* [15] and

conversely, not a single substrate is known for about half of the kinases and phosphatases identified in the same organism.

The work presented in this thesis provides an important step towards the understanding and interpretation of this essential regulatory network in eukaryotes and describes the first system–wide protein phosphorylation network *in vivo*. Our results constitute a major resource for further studies, not only concerning the field of phosphoproteomics but also in a broader Systems Biology context in order to describe the interactions of the various biological networks in living organisms. In doing so, this work further gives insights into the interrelationship of the phosphorylation network with the metabolic network by analyzing the impact of protein phosphorylation on the regulation of the metabolome.

2 PROTEIN PHOSPHORYLATION

Protein phosphorylation is one of the most common post–translational modifications (PTMs) of proteins. As depicted in Figure 1(a), it plays a key role in eukaryotes, regulating essential cellular processes including cell division, signal transduction, metabolism, and cell motility [12, 13]. The process of "enzymatic protein phosphorylation" was first described in 1952 by George Burnett and Eugene Kennedy [16]. The reversible transfer of a phosphate group to specific amino acids is catalyzed by protein kinases while its removal is catalyzed by protein phosphatases (see Figure 1(b)). In eukaryotes, mostly serine and threonine residues are phosphorylated, and, to a lesser extent, also tyrosine residues. Nevertheless, phosphorylation events involving histidine and aspartate residues have also been described [13].

Nowadays, it is common knowledge, that the addition of a phosphate group to the side chain of an amino acid can have strong effects on the characteristics of proteins. Particularly, the protein's charge is changed which in turn can induce extensive conformational modifications and for instance result in the activation or deactivation of enzymes. Also, the protein's ability to interact with other proteins can be influenced. Existing binding sites can be blocked as well as new binding sites created. In consequence, protein phosphorylation is significantly involved in protein complex assembly and disassembly.

It has been estimated that 30 to 50 % of all cellular proteins might be phosphorylated and thus represent possible phosphoproteins [21, 22]. The number of human phosphorylation sites has recently been revised upwards from 100.000 to 500.000, implying an average of about 25 potential sites per protein [23]. In yeast, currently about 9000 phosphosites for over 3000 phosphoproteins are curated [24]. The availability of multiple phosphorylation sites in the same protein (multisite phosphorylation) greatly increases the combinatorial space of regulation and plays a significant factor in the fine–tuned control of protein function. Furthermore, the order in which the sites are processed seems to play a crucial role, too [13, 25].

The kinome and phosphatome represent a species' complete set of protein kinases and phosphatases, respectively. Protein kinases are among the largest of protein families in eukaryotes, comprising about 2% of all eukaryotic genes [22, 26, 12, 27, 28]. For example, the analysis of the human genome revealed the existence of 518 putative protein kinases [12]. In the fruit fly, *Drosophila melanogaster*, 239 protein kinases have been reported [29, 30], in yeast 123 [31]. Based on sequence similarities between their catalytic domains, protein kinases can be classified into a hierarchy of groups,

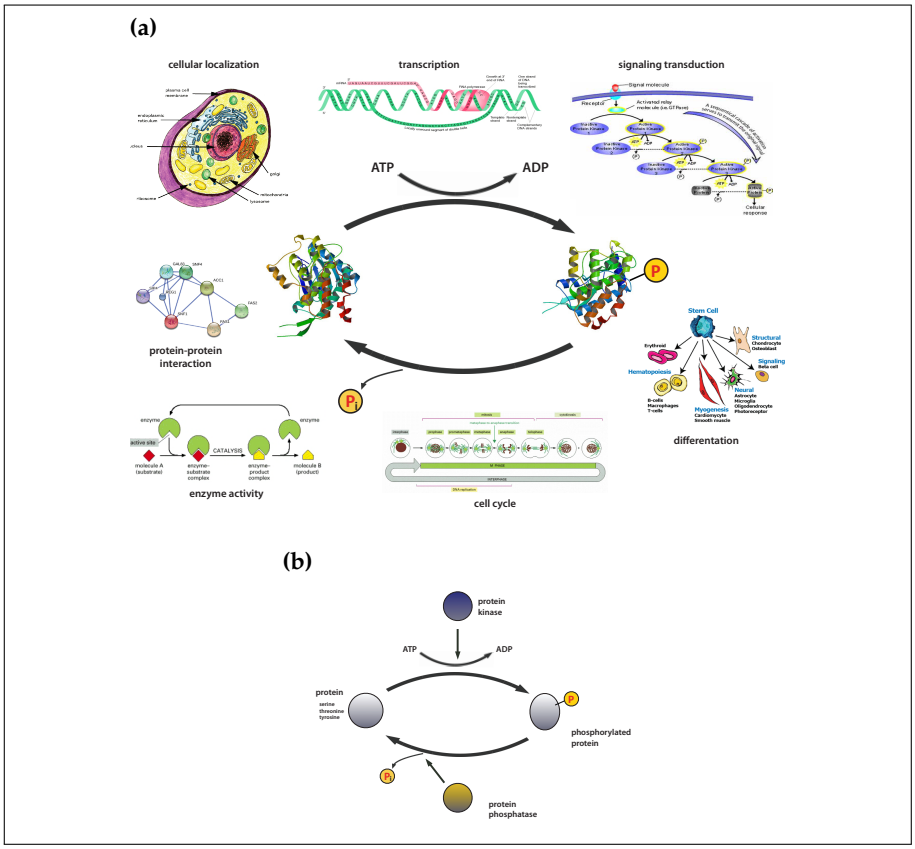


Figure 1: *Reversible protein phosphorylation.* (a) Schematic overview of a selection of cellular processes known to be regulated by protein phosphorylation. Graphics taken from [17, 18, 19, 20] (b) Schematic diagram of reversible protein phosphorylation. In eukaryotes, mostly serine, threonine, and tyrosine residues are phosphorylated. Protein kinases catalyze the transfer of a phosphate group from ATP to an amino acid, protein phosphatases catalyze the removal of a phosphate group.

families, and subfamilies [29, 32]. Phylogenetic comparisons of different eukaryotic kinomes showed an ancient conservation of kinase families: All major kinase groups and most kinase families are shared among metazoans and many of them are also conserved in yeast [29].

As a result on the substantial influence of protein phosphorylation on most cellular processes, abnormal phosphorylation leads to many diseases, including cancer, diabetes, and rheumatoid arthritis [33]. However, this aspect also makes protein kinases and phosphatases to a possible source as drug targets. In fact, kinases have become the second most important group of drug targets and a multitude of human clinical trials involving kinases are carried out at present [34].

3 PHOSPHOPROTEOMICS

The advent of proteomics [35] has revolutionized the way we study proteins. Simultaneously, the analysis of protein phosphorylation has shifted from targeted biochemical studies to high–throughput (HTP) studies, leading to the subdiscipline of phosphoproteomics, dedicated to the systematic study of the phosphoproteome (i.e., the comprehensive set of phosphoproteins in an organism) [36, 37].

Within the proteomics discipline, mass spectrometry (MS) is the technology of choice, and shotgun proteomics (bottom–up proteomics) is the most popular approach used in phosphoproteomics [38, 39]. Figure 2 outlines the standard shotgun proteomics procedure. In a first step, complex protein samples are enzymatically digested into peptides, which are then separated by the MS pipeline. Phosphopeptides are naturally underrepresented in digested peptide mixtures. In order to compensate for this, in phosphoproteomics, protein digestion is followed by the enrichment of phosphopeptides, usually by applying metal affinity chromatography like titanium dioxide (TiO_2) or immobilized metal affinity chromatography (IMAC) [40].

Implementing this approach, hundreds of phosphorylation sites have been identified and reported [42, 43, 44, 45, 46]. Various databases like PhosphoPep [24], Phospho.ELM [23], Phosida [47], and PhosphoSite [48] were established to catalogue and curate these large numbers of phosphosites.

In recent years, analyses have changed from the simple reporting of phosphorylation sites to more quantitative phosphoproteomic studies. This allows obtaining precise measurements of small changes in the level of abundance of phosphopeptides of interest, and hence gaining insights into the dynamics of protein phosphorylation across different cellular states [49, 50, 51, 52].

In principle, there are two main approaches for MS–based quantitative proteomics: isotope label–based quantification and label–free quantification. A number of excellent reviews addressing these methods have been published, for a selection see [53, 54, 39]. Figure 3 gives a schematic overview of both workflows, whereas the MS–landscape in Figure 4 highlights the differences in identifying and quantifying identical peptides in different samples.

Briefly, in isotope labeling proteomics, (Figure 3(a)) peptides are modified metabolically (stable isotope labeling by amino acids in cell culture (SILAC)), chemically (isotope–coded affinity tag (ICAT)), or enzymatically (proteolytic

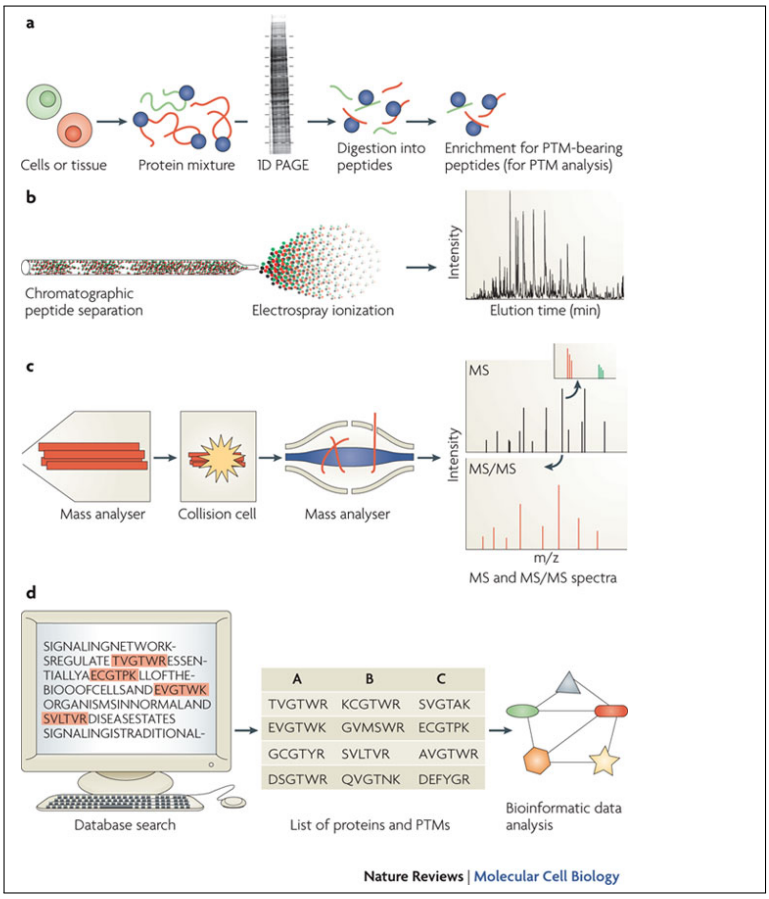


Figure 2: *Standard shotgun proteomics workflow.* (a) Proteins sampled from cells or tissues are digested into peptides and if necessary enriched for peptides containing specific PTMs. (b) Purified peptides are chromatographically separated and then ionized (for example by electrospray ionization (ESI)) directly into the mass spectrometer. (c) In the mass spectrometer particular peptide ions are selected for MS_2 and fragmented in a collision cell, resulting in a MS/MS spectrum. (d) Spectra are mapped to peptides and proteins, which can then be further analyzed via bioinformatic approaches. Taken from [41].

^{18}O labeling). These 'heavy' peptides are then run together with their native counterparts ('light' peptides) and are analyzed by liquid chromatography coupled to tandem mass spectrometry (LC–MS/MS). The quantification is based on the fact that labeled and native peptides are chemically identical and therefore should also behave identically (at least outside the mass spectrometer). As a result, they can be recognized by their (known) mass difference, and by comparing their signal intensities, a change in relative abundance of the two peptides can be accurately determined.

In contrast to labeled proteomics, in label–free approaches different samples are analyzed by separate LC–MS/MS runs (Figure 3(b)). The (relative) quantification is achieved either by aligning peptides across different samples based on their precursor ion scan of the mass spectrometer (MS_1) spectra and followed by a comparison of their peak intensities (comparative LC–MS/MS), or by estimating the abundances of peptides based on their number of identified MS_2 spectra (spectral counting) and comparing the number of such spectra between different samples.

Both labeled and label–free approaches have advantages and disadvantages. In general, the former offer higher accuracy in measuring protein abundances while label–free methodologies exhibit a greater dynamic range and enable more comprehensive (phospho)-proteome coverage [53].

A crucial factor in comparative LC–MS/MS is the computational procedure of matching identical peptides across different samples/MS runs. In the following section, I will explain some of the problems and challenges arising while generating this type of alignment, and review some of the software applications currently available.

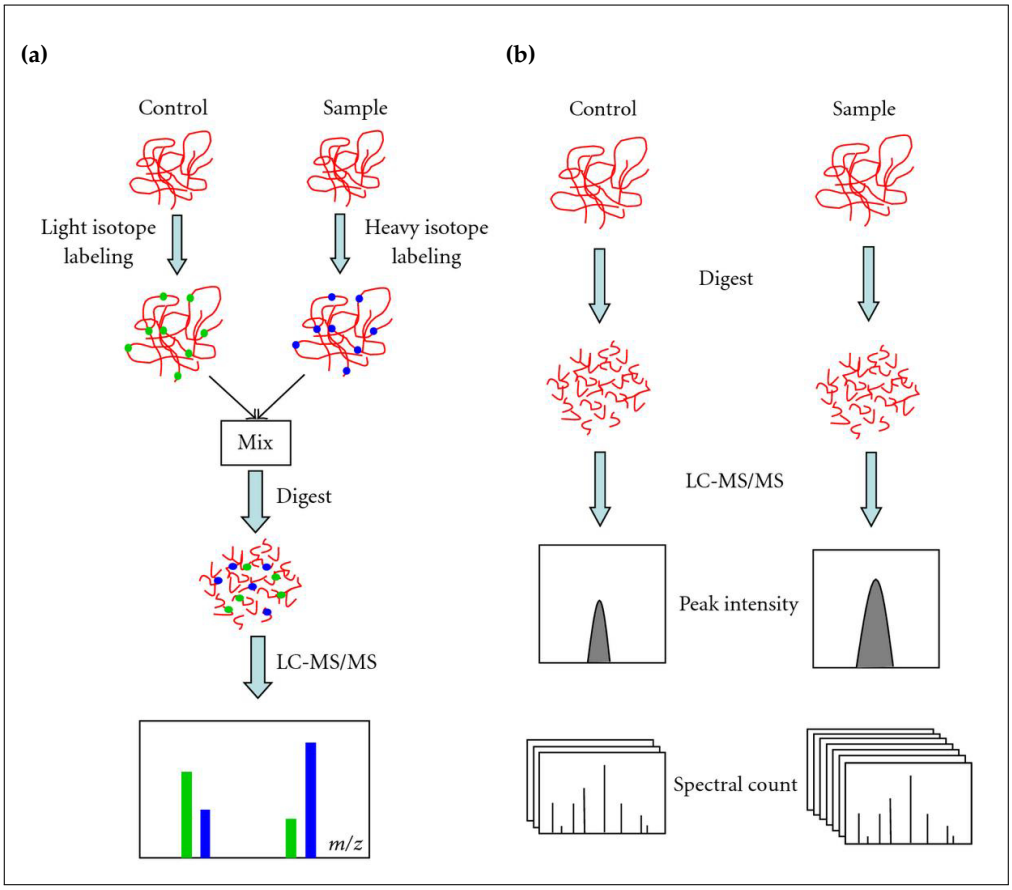


Figure 3: *Main approaches for quantitative proteomics* (a) **Isotope label–based proteomics.** Different samples containing light and heavy isotopes are analyzed together. The quantification is based on the ratio of signal intensities of chemically identical but differently labeled peptides. (b) **Label–free proteomics.** The quantitative information can be achieved either by aligning peak intensities across different samples or by comparing spectral counts, respectively. Both methodologies yield relative protein quantifications. Taken from [55].

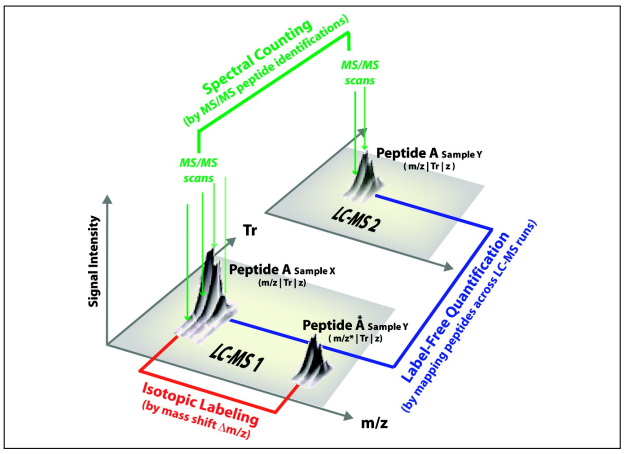


Figure 4: *Comparison of the three main quantification approaches in proteomics.* **Red**: Isotope label–based proteomics. Chemically identical but differently labeled peptides can be identified based on their mass shift. **Blue**: Comparative LC–MS/MS. Identical peptides in different samples are identified by their specific retention time and precise m/z values. **Green**: Spectral counting. Identical peptides are identified based on their MS_2 annotation. Taken from [56].

4

COMPARATIVE LC–MS/MS

Label–free quantification based on the comparison of precursor ion intensities depends on the accurate alignment of corresponding MS1 spectra across the different liquid chromatography mass spectrometry (LC–MS) runs and requires prior peak and feature detection. Typically, most software programs for comparative LC–MS/MS do implement these steps. A brief outline of the usual data processing pipeline of label–free quantification software is depicted in Figure 5.

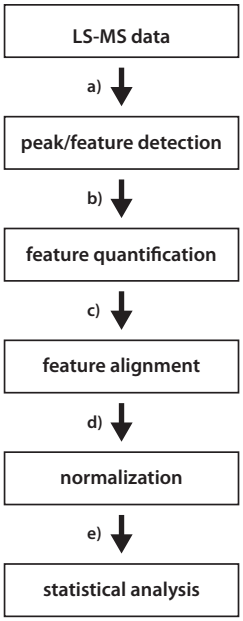


Figure 5: *Typcial data processing workflow of comparative LC–MS/MS software*. **a)** Raw MS1 data obtained from the mass spectrometer is processed for feature detection, resulting in a list of features characterized by their m/z and Rt values. **b)** Features are quantified by the integration of their signal intensities. **c)** Mass spectra of different MS runs are aligned and identical features are matched. **d)** Usually, this is followed by the normalization of Rt and intensity values (in some algorithms this is done at the beginning of the pipeline). **e)** Statistical analysis: identification of features that differ significantly in their abundances across the different MS runs.

Peak detection is the process of selecting peptide–induced signals from MS raw data. In doing so, one of the major bottlenecks is distinguishing between peptide and noise signals. Both chemical noise (molecules other than peptides) and random (electronic) noise contribute to the overall noise rate. Additionally, the true peptide signal is disturbed by the baseline, which is the background signal produced by the mass spectrometer. Consequently, noise filtering or 'smoothing' and baseline correction mark the beginning of peak detection algorithms. For smoothing, most approaches are based on signal processing techniques like Gaussian or wavelet–based filters. Baseline correction is done by subtracting the estimated baseline signal from the

peptide signals. Reference [57] summarizes some of the standard smoothing filters and baseline modulation methods.

The actual peak extraction procedure ('peak picking') utilizes a set of criteria like signal–to–noise (S/N) ratio, miscellaneous peak characteristics (shape, intensity, slope, width), and the existence of local maxima [57]. Due to the occurrence of isotopes on the one hand and different charge states on the other hand, peptides are never represented as single peaks but as a sequence of isotopic peak patterns. Thus, in a last step all peaks arising from the same peptide are grouped and converted into a single feature. Often, only the peak representing the monoisotopic mass (i.e. the mass calculated while taking only the most abundant isotope of each constituent atom into account) is used for later quantification. The separation of overlapping peak patterns constitutes a complex matter and is a likely source of error.

The quantification of features is based on the finding that the abundance of peptides in a sample should be proportional to their MS peak intensity [58]. Therefore, quantification is achieved by integrating the peak intensity over chromatographic time (also referred to as area under the curve (AUC)). Because the final goal of the analysis is to compare peptide abundances across different samples, it is indispensable to ensure identical experimental and instrumental conditions for all the protein samples analyzed sequentially.

The main task in comparative LC–MS/MS is constructing the alignment of multiple LC–MS runs in order to compare peptide quantities across different samples. Peaks identified in one sample are to be linked to matching peaks in the other sample. This constitutes a nontrivial challenge, and, as stated in [58], "given the enormous complexity of high resolution LC–MS datasets, a perfect detailed alignment of all features seems a non–realistic goal."

Below is a listing of challenges that algorithms have to deal with:

- drifts in Rt and m/z between different runs, caused by technical problems
- Rt and m/z drifts often differ within the set of identified peaks (mainly non–linear shifts in Rt)
- a single peak in one run could be interpreted as multiple peaks in the other run (inaccurate peak detection)
- peaks could be missing completely in one sample, either due to true absence or failure of detection
- features having very similar Rt and m/z values can nevertheless represent different peptides (high–complexity peptide samples)

Though comparative LC–MS/MS is a relatively recent field, a variety of different software solutions is available in order to master the challenge of

LC–MS alignment construction. All of them have in common that the quality of the alignment built is highly dependent on the resolution of the MS data in terms of mass accuracy and chromatographic separation. The higher the resolution in the m/z dimension, the better peaks will be discriminated.

Well–established software applications for LC–MS alignment construction include SuperHirn [59], OpenMS [60], msInspect [61], SpecArray [62] and MZmine [63], among others. Recently, MaxQuant, a frequently used software in the field of isotope–label based proteomics analysis [64], is available with a label–free implementation, too [65].

5

PREVIOUS WORK ON THE RECONSTRUCTION OF PHOSPHORYLATION NETWORKS

Deciphering (global) interrelationships between kinases, phosphatases and their substrates has always been the main focus of protein phosphorylation analysis. In order to link kinases and phosphatases to their substrates, besides MS–based phosphoproteomics other HTP techniques have been employed as well. In the following paragraphs, I outline some of the relevant research studies of recent years. Because the work in my thesis is done using yeast as a model organism, I will focus on findings in this system.

One of the first HTP technologies employed were *in vitro* phosphorylation assays. Based on them, Ubersax et. al [66] identified about 180 targets of yeast Cdk1. In a similar study, 24 substrates of yeast kinase Pho85 could be determined [67].

Ptacek et al. used proteome microarrays to determine *in vitro* substrates of yeast protein kinases [68]. For 87 kinases they identified about 4200 phosphorylation events comprising more than 1300 phosphoproteins. Most of the targets were recognized by only one or two kinases, indicating a strong preference of kinases for specific targets. By integrating protein protein interaction (PPI) and transcription factor (TF) binding site data, they constructed the first integrated *in vitro* phosphorylation network in yeast.

Detection of functional connections between kinases and substrates by applying genetic screens is discussed in [69]. Comprising most of the known kinases and phosphatases in yeast, the authors measured about 100.000 distinct pairwise genetic interactions. Their analysis revealed a significant enrichment of positive genetic interactions between known kinase–substrate and phosphatase–substrate pairs. Particularly strong positive genetic interactions could be observed for kinase–phosphatase pairs that shared a common substrate. Therefore, the authors argued that their dataset would be helpful in identifying unknown kinase/phosphatase substrates.

Combining affinity purification and MS, Breitkreutz et al. [70] constructed a global kinase and phosphatase interaction network in yeast, consisting of more than 1800 interactions between almost 900 proteins. A clustering analysis revealed many locally dense regions, suggesting new functions for known kinases and phosphatases. Upon the incorporation of (previously known) PPI datasets, an extended interaction network showed a significant enrichment in kinase–kinase interactions, strongly suggestion that the extensive kinase–kinase interaction network constitutes the proteome backbone.

Van Wageningen et al. integrated genome–wide transcriptomic profiling and genetic screens to study the degree of redundancy between kinases and phosphatases in signaling networks [71]. Three different genetic buffering relationships were identified: mixed epistasis, complete redundancy, and quantitative redundancy. Mixed epistasis, the most frequent type of redundancy detected, is exhibited by pairs of kinases and phosphatases that show only partial functional overlaps but that are coupled by additional regulatory links. The authors conclude that such modules act as the molecular mechanism for multi–process control and that they likely contribute to the stable maintenance of paralogs during evolution.

A computational approach to determine kinase–substrate relationships is presented in [72]. Here, the software NetworKIN combines sequence models of kinase consensus motifs with contextual information about putative targets in order to predict *in vivo* kinases for identified phosphorylation sites. Context information comprises and integrates a) experimental evidence data like physical interactions and co–expression, b) genomic context like gene fusion, gene neighborhood, and phylogenetic profiles information as well as c) manually curated pathway databases, and d) automated literature mining. In a thoroughly investigation, the authors could show that NetworKIN improves the accuracy of phosphorylation network prediction by 2.5 fold compared to previous methods, making it a useful tool for the construction of phosphorylation networks.

Recently, a database for literature–curated kinase–substrate pairs was introduced [73]. KID, the Yeast Kinase Interaction Database, holds both HTP as well as low–throughput (LTP) data, adopted from various sources. By integrating these datasets quantitatively, a set of 517 high-confident kinase–substrate pairs was defined and suggested to act as a gold standard in order to assess published HTP datasets.

By now, the availability of phosphoproteome datasets in different organisms provides the opportunity to study protein phosphorylation in the light of evolution. Below I summarize some of the research studies addressing this issue.

Boekhorst and co–workers conducted one of the first comparative analysis of protein phosphorylation based on public available phosphoproteomics data of distantly related species [74]. In their research they compared the phosphoproteomes of six eukaryotes (*Homo sapiens*, *Mus musculus*, *Drosophila melanogaster*, *Arabidopsis thaliana*, and *Danio rerio*). They found the overlap between the phosphoproteomes to be significantly enriched, indicating an increased functional relevance. However, due to the incomprehensive nature of HTP datasets and differences in experimental procedures, they clearly pointed out the need of additional datasets in order to improve the power of comparative phosphoproteomics to understand protein phosphorylation evolution.

In order to gain insights into the evolution of phosphorylation sites, Landry et al. compared high quality phosphoproteomics datasets of yeast and human [75]. Interestingly, their analysis revealed differences in the evolution of phosphosites in ordered and in disordered regions, as well as of phosphosites with known function and phosphosites that are uncharacterized. Phosphosites in ordered regions exhibit a higher degree of conservation, and the same observation holds true for functional phosphosites, indicating that the known rapid evolution of phosphoproteomes is possibly due to a large set of non–functional phosphosites.

One of the first system–wide studies to quantify evolutionary changes in phosphoproteomes across different species was conducted by Beltrao et al. [76]. They applied *in vivo* MS to compare the phosphorylation status of proteins in *Saccharomyces cerevisiae*, *Candida albicans*, and *Schizosaccharomyces pombe*, and estimated the evolutionary rate at which changes in phosphorylation emerged. Their analysis revealed that interactions between kinases and their substrates evolve at most two orders of magnitude slower than those between TFs and deoxyribonucleic acid (DNA) promoter sequences. Furthermore, the incorporation of genetic interaction data showed that genetic interactions among kinases are less conserved than expected, suggesting a substantial contribution of kinases to the evolution of phenotypic diversity.

Tan et. al studied the conservation of human phosphorylation sites in *Saccharomyces cerevisiae*, *Drosophila melanogaster*, and *Caenorhabditis elegans* by means of a comparative sequence–alignment analysis [77]. They found about 480 human phosphorylation events that exhibit an ancient conservation in their position. Additional 780 conserved phosphorylation sites could be identified by an integrated network–alignment approach. Interestingly, the proteins corresponding to conserved phosphosites were found to be enriched in disease–associated genes, suggesting that this type of analysis could be helpful to identify networks that are misregulated in diseases like cancer or diabetes. Furthermore, the authors argue that conservation of kinase–substrate interactions is to be considered equally important as sole conservation of phosphoproteins when searching for disease–related genes.

Another approach to study the evolutionary properties of phosphorylation sites is presented in [78]. Here, annotated phosphorylation sites from *Saccharomyces cerevisiae* were compared to their respective flanking regions in closely related species of yeasts (*Saccharomyces paradoxus*, *Saccharomyces mikatae*, and *Saccharomyces bayanus*). As a result, not only phosphorylation sites but also the phosphorylation motifs recognized by the kinases are significantly higher conserved than the surrounding sequences, indicating that their evolution is constrained relative to the regions in which they occur.

In contrast to the work described above, the research presented in [79] did not reveal a particular conservation in the position of most phospho-

rylation sites analyzed. Substrates of yeast kinase Cdk1 were identified by quantitative MS and compared to orthologous sequences of 32 fungal species. The results indicated a conservation of phosphosites in disordered regions but less constraints regarding the precise positions of the sites. More precisely, the conservation appeared only to be conserved in very closely related species.

A very recent study addresses the evolution of phosphorylation networks in the context of gene duplication [80]. The authors compared the phosphorylation sites of paralogous proteins in yeast to a species that diverged from yeast prior the duplication of those genes. Their analysis revealed a considerable divergence of paralogous phosphoproteins in terms of their sequence, function, localization and/or recognition by protein kinases, meaning that only only 36 – 54% of regulation by phosphorylation seems to be conserved, even if a conservation of the actual residue is known.

Part II

THE PHOSPHORYLATION NETWORK OF *SACCHAROMYCES CEREVISIAE*

6

PREFACE

Phosphorylation networks describing the interdependences between kinases and phosphatases and their substrates constitute highly dynamic systems. The process of reversible protein phosphorylation affects most cellular activities including cell division, signal transduction, metabolism and cell motility. In recent years, large-scale MS studies provided growing amounts of phosphoproteomic data, characterizing phosphoproteins and their sites of phosphorylation in a variety of organisms. Nevertheless, for the majority of phosphorylation sites their signaling context is unknown, and linking the kinome and phosphatome (the complement of kinases and phosphatases, respectively, in an organism) to the phosphoproteome remains a challenging task.

In our study, we analyzed kinase and phosphatase signaling in the yeast *Saccharomyces cerevisiae*. Based on perturbed phosphoproteome data, we identified phosphopeptides whose abundances reproducibly changed upon the absence of a given kinase or phosphatase. Thereby we generated a network describing the first system–wide *in vivo* protein phosphorylation network of any eukaryotic organism.

For this study, I developed and carried out the computational pipeline to compile and analyze the phosphoproteomic data in order to construct the yeast protein phosphorylation network. The experiments that generated the raw data were performed by Bernd Bodenmiller, ETH Zurich.

In detail, I implemented the clustering algorithm for the subsequent processing of the LC–MS/MS alignments produced by SuperHirn [59] (for details, see Supplementary Materials). This post–alignment procedure compensates errors introduced during peak detection and peak matching and ensures high–quality LC–MS/MS alignments between the wild–type and perturbed phosphoproteomes.

Furthermore, I assisted in the statistical analysis of the data, both by using Corra [81] and by employing a self coded customized version of the Bioconductor Limma package [82].

Particularly, I implemented the bioinformatics workflow for the functional analysis of the data. In addition, I compiled and prepared various phosphoproteomic, transcriptional, genomic interaction and PPI datasets in order to compare and integrate them with our data.

Last but not least, I contributed to most of the figures. In particular figures 1–3, and figures S2–S7.

7

PHOSPHOPROTEOMIC ANALYSIS REVEALS INTERCONNECTED SYSTEM–WIDE RESPONSES TO PERTURBATIONS OF KINASES AND PHOSPHATASES IN YEAST

The original publication is included below and is followed by a selection of the Supplementary Materials, featuring my contributions.

SYSTEMS BIOLOGY

Phosphoproteomic Analysis Reveals Interconnected System-Wide Responses to Perturbations of Kinases and Phosphatases in Yeast

Bernd Bodenmiller,[1,2*†] Stefanie Wanka,[2,3*] Claudine Kraft,[4] Jörg Urban,[5] David Campbell,[6] Patrick G. Pedrioli,[4‡] Bertran Gerrits,[7§] Paola Picotti,[1] Henry Lam,[8] Olga Vitek,[9] Mi-Youn Brusniak,[6] Bernd Roschitzki,[7] Chao Zhang,[10] Kevan M. Shokat,[10] Ralph Schlapbach,[7] Alejandro Colman-Lerner,[11] Garry P. Nolan,[12] Alexey I. Nesvizhskii,[13] Matthias Peter,[4] Robbie Loewith,[5] Christian von Mering,[3] Ruedi Aebersold[1,6,14‖]

(Published 21 December 2010; Volume 3 Issue 153 rs4)

The phosphorylation and dephosphorylation of proteins by kinases and phosphatases constitute an essential regulatory network in eukaryotic cells. This network supports the flow of information from sensors through signaling systems to effector molecules, and ultimately drives the phenotype and function of cells, tissues, and organisms. Dysregulation of this process has severe consequences and is one of the main factors in the emergence and progression of diseases, including cancer. Thus, major efforts have been invested in developing specific inhibitors that modulate the activity of individual kinases or phosphatases; however, it has been difficult to assess how such pharmacological interventions would affect the cellular signaling network as a whole. Here, we used label-free, quantitative phosphoproteomics in a systematically perturbed model organism (*Saccharomyces cerevisiae*) to determine the relationships between 97 kinases, 27 phosphatases, and more than 1000 phosphoproteins. We identified 8814 regulated phosphorylation events, describing the first system-wide protein phosphorylation network in vivo. Our results show that, at steady state, inactivation of most kinases and phosphatases affected large parts of the phosphorylation-modulated signal transduction machinery, and not only the immediate downstream targets. The observed cellular growth phenotype was often well maintained despite the perturbations, arguing for considerable robustness in the system. Our results serve to constrain future models of cellular signaling and reinforce the idea that simple linear representations of signaling pathways might be insufficient for drug development and for describing organismal homeostasis.

INTRODUCTION

Protein kinases, and, to a lesser extent, protein phosphatases, are attractive drug targets (*1*–*5*); however, although their respective catalytic activities are well characterized, their functions in vivo remain relatively poorly understood. Despite extensive in vitro (*6*), in silico (*7*), or indirect in vivo assays (*8*), our knowledge of the global relationships between kinases, phosphatases, and their substrates remains fragmented (*2*). Even less is known about the more downstream, indirect consequences of kinase activity, making rational selection of suitable candidates for therapeutic interventions difficult; consequently, many promising kinase inhibitors are ultimately retired from development (*9*).

One promising approach for closing this knowledge gap is the organism-wide, quantitative assessment of all phosphorylated proteins, comparing phosphorylation status in wild-type cells to that in cells that have undergone systematic perturbations of their kinases or phosphatases. Progress in phosphoproteomics technology has brought this goal within reach by enabling the reproducible quantification of thousands of phosphorylation sites in a single study (*10*–*12*). Although the throughput is not yet sufficient to systematically address all 518 protein kinases and 147 protein phosphatases in human cells (*13*, *14*), simpler organisms, such as yeast, can be addressed. Yeast in particular is frequently used as a model to study human diseases (*15*), including cancer, mitochondrial diseases, and even neurological disorders caused by protein misfolding (*16*, *17*).

[1]Institute of Molecular Systems Biology, ETH Zurich, 8093 Zurich, Switzerland. [2]Zurich PhD Program in Molecular Life Sciences, 8057 Zurich, Switzerland. [3]Institute of Molecular Life Sciences and Swiss Institute of Bioinformatics, University of Zurich, 8057 Zurich, Switzerland. [4]Institute of Biochemistry, ETH Zurich, 8093 Zurich, Switzerland. [5]Department of Molecular Biology, University of Geneva, Geneva 1211, Switzerland. [6]Institute for Systems Biology, Seattle, WA 98103, USA. [7]Functional Genomics Center Zurich, University Zurich and ETH Zurich, 8057 Zurich, Switzerland. [8]Department of Chemical and Biomolecular Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong. [9]Departments of Statistics and Computer Science, Purdue University, West Lafayette, IN 47107, USA. [10]Howard Hughes Medical Institute and Department of Cellular and Molecular Pharmacology, University of California, San Francisco, CA 94158–2280, USA. [11]Facultad de Ciencias Exactas y Naturales, University of Buenos Aires, C1428EHA Buenos Aires, Argentina. [12]Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, CA 94305, USA. [13]Department of Pathology, University of Michigan, Ann Arbor, MI 48109, USA. [14]Faculty of Science, University of Zurich, 8057 Zurich, Switzerland.
*These authors contributed equally to this work.
†Present address: Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, CA 94305, USA.
‡Present address: Scottish Institute for Cell Signalling, Sir James Black Centre, University of Dundee, Dundee, Scotland DD1 5EH, UK.
§Present address: Novartis Institute for Biomedical Research, Novartis International AG, CH-4002 Basel, Switzerland.
‖To whom correspondence should be addressed. E-mail: aebersold@imsb.biol.ethz.ch

Although some signaling systems, such as the apoptotic machinery, are absent in yeast, other parts of its signaling network display substantial similarities to those in human cells (*18*, *19*). Of the 161 kinases and phosphatases in yeast, 136 are conserved in humans at more than 30% amino acid sequence identity (table S1), and some human signaling proteins can even replace their yeast counterparts (*20*). Here, we used a combination of phosphoproteomics measurements and computational methods (*11*) to detect and quantify the system-wide responses in the yeast phosphoproteome upon deletion or inhibition of most of its kinases and phosphatases.

RESULTS

Experimental strategy

We developed an integrated experimental and computational strategy for high-throughput comparative phosphoproteomic analysis in *Saccharomyces cerevisiae* (Fig. 1), which consisted of the following steps. First, we systematically perturbed the kinase-substrate and phosphatase-substrate networks by selecting gene deletion mutants of the nonessential kinases or phosphatases or, for some essential kinases, by generating mutants inhibitable by cell-permeable drugs, which are referred to as "analog-sensitive" kinase strains (*21*). To minimize compensatory mutations that might accumulate over time in the gene deletion strains, we freshly prepared all mutant strains. To enable a statistical characterization of our observations, we always grew, processed, and measured each perturbed strain in three independent replicates, together with three replicates of wild-type, control cells. Phosphopeptides were isolated from each sample (*22*, *23*) and submitted to high-performance mass spectrometry to generate liquid chromatography coupled to mass spectrometry LC-MS/MS phosphoproteome maps. The triplicate phosphoproteome maps generated from each perturbed or wild-type cell sample were annotated with the amino acid sequences of the detected phosphopeptide features and were aligned with the algorithm SuperHirn (*24*), which was followed by additional postprocessing (see Supplementary Materials for details). The statistical significance of observed changes in the perturbed states was then computed for each phosphopeptide with the Corra software suite (*25*).

We assessed the reliability of our measurements and computational data processing at two levels. First, we assessed the confidence of the phosphopeptide identifications generated by database searching, and second, we assessed the reproducibility of detecting quantitative phosphopeptide differences between wild-type and mutant strains. For the first check, and to determine the reliability of our phosphopeptide identifications from the peptide fragment ion spectra, we performed statistical analyses with the PeptideProphet tool (*26*) and a decoy database strategy (*27*). From these analyses, we found that a PeptideProphet probability cutoff of 0.9 corresponded to a false discovery rate (FDR) of ~0.038 (3.8%) (table S2), which confirms that our chosen cutoff of 0.9 yielded an acceptably low degree of incorrect peptide identifications, in particular because most phosphopeptides were identified repeatedly in the context of this extensive study.

We then used the statistical tool Corra (*25*), which supports an empirical Bayesian alternative to the *t* test (*28*). The test improves the reliability of conclusions in cases of large-scale testing. For each phosphopeptide feature, the test provided a *P* value of the observed differences between wild-type and mutant replicates. The *P* values were further corrected for multiple testing according to the Benjamini and Hochberg procedure (*29*) (see the Supplementary Materials). After this quantitative analysis step, we chose an FDR threshold of 0.015 in conjunction with a minimum fold-change requirement of $\log_2 > 1.5$, both of which had to be met before we would consider any phosphopeptide as reproducibly regulated. At this threshold, nine comparisons between wild-type and lowest-impact kinase mutants resulted in only a single or no phosphopeptide being designated as regulated, which verified the validity of our selected criteria. On the basis of these results, we concluded that our applied cutoffs ensured that, despite a high sensitivity (fig. S1), only a minimal amount of noise entered our analyses and that we achieved high reproducibility in the observed regulatory events.

Overall, we attempted the analysis of 161 mutant strains of yeast. Of these, 37 strains could not be analyzed because they were not viable, not inhibitable, or otherwise not amenable to our procedure (table S1). In total, we generated quantitative data for 116 gene deletion mutants and for an additional 8 strains in which analog-sensitive kinases were pharmacologically inhibited (table S1). Together, this corresponds to coverage of 78% of the theoretical kinase and phosphatase space in yeast and covered

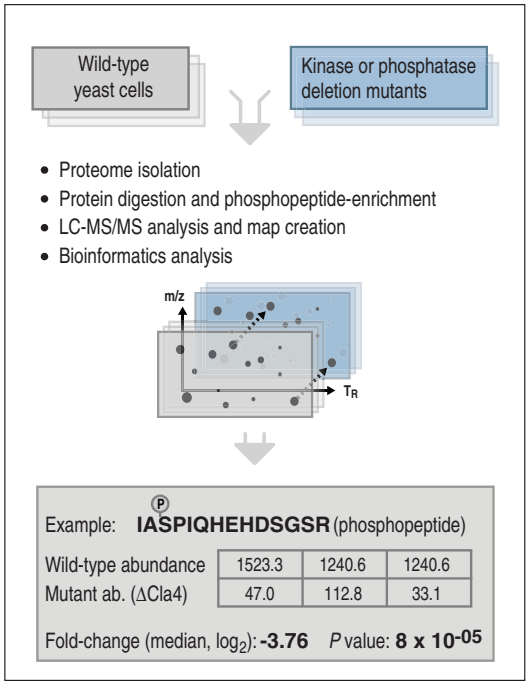


Fig. 1. Integrated experimental and computational pipeline to determine in vivo kinase-substrate and phosphatase-substrate relationships. Yeast kinase and phosphatase genes were systematically deleted one by one and the phosphoproteomes were systematically compared between mutant and wild-type strains. To achieve this, for each mutant strain and wild-type, we grew and processed three independent biological replicates by proteome isolation, protein digestion with trypsin, phosphopeptide enrichment by applying a TiO_2 resin, and quantification and identification of the phosphopeptides with LC-MS/MS. Observed phosphopeptide ion features were aligned, quantified, and tested for statistical significance. For the example phosphopeptide shown, IAS*PIQHEHDSGSR, the resulting matrix gives the intensity values measured in the wild-type and mutant samples, as well as the corresponding $\log_2$ fold change (here –3.76) with its associated significance. Abbreviations for the amino acids are as follows: A, Ala; D, Asp; E, Glu; G, Gly; H, His; I, Ile; P, Pro; Q, Gln; R, Arg; and S, Ser.

77% of those enzymes that show sequence conservation with human enzymes (table S1). A matrix and a network generated from these data related the observed changes in the abundance of a phosphopeptide (measured in triplicate) to the corresponding kinase or phosphatase deletion (Fig. 2 and fig. S2). The matrix contains 8814 reproducible changes in peptide abundance that mapped to 1026 phosphoproteins that were clustered according to the coregulation of the phosphopeptides (tables S3 and S4). Of note, an additional 7550 phosphopeptides were consistently identified but did not exhibit a substantial change in abundance under any of the perturbations tested.

Finally, the cellular abundance distribution of detected phosphoproteins (regulated and unregulated) was roughly similar to that of the total yeast proteome; however, the complete phosphoproteome was still not covered (fig. S3), because under our chosen growth conditions, many phosphorylation sites would not be phosphorylated, and because our experimental pipeline had several biases, among them that only tryptic peptides with a mass/charge ratio (m/z) suitable for LC-MS/MS analysis (*30*) could be identified. Nevertheless, the observed phosphorylation sites covered a reasonably large fraction of the phosphoproteome, and therefore an existing bias should not impair our conclusions (*31*).

Direct versus indirect phosphorylation events

Because kinases and phosphatases are components of complex, interconnected signaling networks, we fully expected to observe a number of indirect, downstream responses, that is, phosphopeptides whose abundance would change despite their not being a direct molecular target of the kinase or phosphatase in question. Indeed, we found that such events seemed to strongly outnumber direct kinase-substrate interactions, as argued by the following observations. First, we determined for each kinase or phosphatase the number of phosphopeptides whose responses showed the expected directionality (that is, reductions in abundance in the case of kinase deletions and increases in abundance in the case of phosphatase deletions). In general, the number of phosphopeptides that responded in the expected directionality was roughly similar to that of phosphopeptides that responded with "inverted" directionality (Fig. 2 and fig. S4). Exceptions to this finding were analog-sensitive kinases that were inhibited over the short term; for example, in the case of Cdc28, about 76% of the phosphopeptides were regulated in the expected directionality. No difference in the direction of regulation was observed between nonessential kinases or phosphatases (fig. S4). Second, we conservatively assumed that phosphopeptides that changed in abundance in only a single deletion strain might be direct molecular targets of the kinase or phosphatase in question. By this measure, we found that, at most, 32% of the observed regulatory events might have been direct for kinases (that is, that the events mapped to just a single kinase), whereas in the case of phosphatases this number was 53%. The data sets generated by the short-term inhibition of the analog-sensitive kinases showed a higher fraction of potential direct targets (44%) than did the permanent deletion strains.

Third, we tested the overlap of our data with various previously established reference protein-protein interactions in yeast (*32*–*35*), such

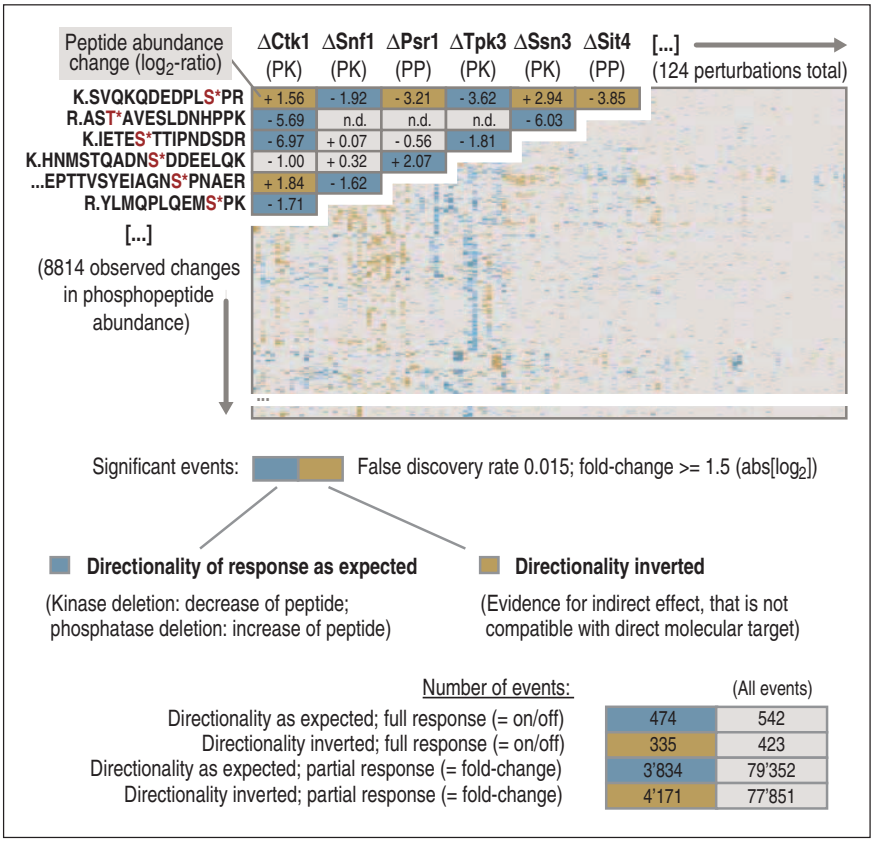


Fig. 2. Matrix of kinases and phosphatases analyzed in this study and their effects on the phosphoproteome. Overall, 124 kinases and phosphatases were interrogated through our experimental and computational pipeline. Each row (*y* axis) corresponds to a regulated phosphopeptide and each column (*x* axis) summarizes the responders of a given kinase or phosphatase. Phosphopeptides with a directionality as expected (that is, kinase deletion resulted in a decrease in peptide abundance, whereas phosphatase deletion resulted in an increase in peptide abundance) are shown in graded blue, and phosphopeptides with an inverted directionality (evidence for indirect effect, not compatible with direct molecular target) are displayed in graded gold, according to the observed fold change for each peptide. Phosphopeptides observed but not regulated or not detected are displayed in gray. At the bottom, the total numbers of events observed in this study are listed. "Full response" corresponds to phosphopeptides that appeared or vanished when wild-type and mutant strains were compared, and "partial response" corresponds to phosphopeptides that showed a statistically significant change in abundance, but were detected in both wild-type and mutant samples. Abbreviations for the amino acids are as follows: A, Ala; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; and V, Val.

as the STRING database (tables S5 and S6). We observed that the overlap of our data with these direct interactions was small (table S5). This is consistent with the long-held notion that kinase-substrate interactions are too weak and transient to be detectable by typical affinity purification–based protein interaction screens. Reassuringly, however, first, the overlap of the heavily studied kinase Cdc28 with our data set on the level of regulated phosphoproteins was high, showing a 43% overlap with the study of Ubersax *et al*. (*36*) and a 76% overlap with the study of Holt *et al*. (*10*) (on the phosphorylation site level, the overlap was 46%). Second, all other phosphorylation events that did overlap showed substantial enrichments for the expected directionality. Likewise, we observed substantial enrichment of confirmed interactions, in particular for those phosphopeptides that responded only in a single perturbation (table S7). This indicates that our data included a sizeable fraction of direct enzyme-target interactions; however, from all three tests, we can conclude that indeed a large majority of our observed events were indirect consequences of the deletion. Not a single kinase showed exclusively direct effects, indicating that a focused modulation of a pathway (branch) without system-wide adaptations might not be possible with a single drug.

Changed extents of phosphorylation versus changed protein abundance

As is the case in prolonged pharmacological intervention, our genetic kinase-deletion approach gave the cells ample time to accommodate (and potentially compensate for) the loss of kinase activity. This should not only have led to downstream, indirect consequences on the phosphoproteome, but could have also entailed subsequent changes in gene expression and the amounts of proteins produced. To assess the extent of this effect, we measured not only abundance changes in the phosphoproteome but also abundance changes of the proteins themselves, by observing unphosphorylated peptides in a subset of 16 kinase deletion strains. The kinases selected for this test ranged from those that had a small effect on the phosphoproteome to those that had a large effect. The data indicated that for a total of 467 regulated phosphopeptides that matched to 118 proteins covered in this analysis, 79% of the proteins remained unchanged in abundance, and, in a single case, the directionality of the phosphopeptide regulation was opposite to the protein abundance change (figs. S5 and S6). In 21% of the cases in which a phosphopeptide was regulated, we also observed a change in protein abundance in the same direction.

We also performed additional orthogonal, but more indirect, analyses based on the coregulation or antiregulation of phosphorylation sites on the same protein, which we found in more than half of the phosphoproteins. We reasoned that a synchronous change with a similar amplitude and directionality of such phosphopeptides would indicate an abundance change of the corresponding protein. In contrast, a discordant abundance change of the phosphopeptides from such proteins would indicate a change in phosphorylation site occupancy. These data (fig. S7) can be summarized as follows: For about 25% of the observed events, only a single regulated phosphopeptide was detected on the entire length of the phosphoprotein, impeding this type of analysis. The remainder of events fell into three classes: In 49% of the remaining cases, at least two phosphopeptides originating from the same protein were observed to be regulated, and these exhibited identical directionality. In contrast, in 5% of events, the changes were of opposing directionality; the latter pattern was not consistent with a simple protein abundance change. Of note, in a large part of the data, that is, in 46% of cases, a phosphopeptide that had substantially changed in abundance was detected with at least one other phosphopeptide on the same protein, but the other phosphopeptides were not observed to be regulated. The latter two categories indicate that for most events detected in this study, changes in the abundance of a phosphopeptide could not be explained by changes in protein abundance alone.

Effect of a given kinase or phosphatase on the phosphoproteome

The number of phosphopeptides that were affected by the deletion of a given kinase or phosphatase varied considerably (Fig. 2). Therefore, we (i) quantified the impact of each kinase or phosphatase on the phosphoproteome under the growth conditions tested, (ii) assessed whether the kinases and phosphatases were associated with different biological processes according to their effect on the phosphoproteome, and (iii) determined which biological processes were affected by each kinase and phosphatase.

We first computed the fraction of phosphopeptides that were affected by a given kinase or phosphatase relative to the total number of phosphopeptides that were affected by the kinases and phosphatases (Fig. 3A and table S8). We observed that the deletion of 22% of the kinases and phosphatases that we tested resulted in fewer than 10 perturbed phosphopeptides each; therefore, we considered these deletions to have had minimal effects on the fraction of the phosphoproteome detected in this study. These included kinases important in cellular stress response mechanisms, such as Mrk1 (*37*) and Gcn2 (*38*). In contrast, for 78% of the kinase and phosphatase deletion strains, distinct changes in the phosphoproteome could be detected. The kinases with the largest effects on the phosphoproteome were Ctk1 (*39*), a kinase with key roles in the regulation of transcription and translation, and Psk2, which is involved in sugar flux and translational regulation (*40*). These data show that the loss of most kinases or phosphatases indeed perturbed large parts of the signaling network.

We next determined the distribution of biological processes represented by the phosphoproteins affected by the lower-impact (bottom half) and higher-impact (top half) kinases and phosphatases, respectively. We found that the enzymes with the smallest effect showed a strong enrichment in processes associated with mitogen-activated protein kinase (MAPK) cascade signaling ["MAPKKK (MAPK kinase kinase) cascade," $P = 3.9^{-10}$; "response to pheromone," $P = 4.2^{-6}$], whereas the enzymes with the largest effects showed a strong enrichment in processes related to the mitotic cell cycle ("interphase of mitotic cell cycle," $P = 3.1^{-9}$; "mitotic cell cycle," $P = 1.4^{-6}$) (tables S9 and S10). These data showed that under the tested conditions, even stress- or mating-related kinases showed a measurable impact on the phosphoproteome, albeit lower than that of growth- and cell cycle–related kinases or phosphatases. Lastly, we also computed those biological processes that were enriched among the responders of each individual kinase or phosphatase. We found that 575 biological processes were enriched (Fig. 3B and table S11), an average of five processes for each active kinase or phosphatase. The most frequently enriched functions were "endocytosis" (39 times) and "cell morphogenesis" (38 times). Together, these data illustrate that the effects of most kinases and phosphatases on the signal transduction network, and thereby on controlled biological processes, were broad, perhaps broader than expected (*2*).

Correlation with yeast phenotypes

We next tested the phenotypic consequences of deletion of kinases and phosphatases, which are relevant in particular with regard to effects (side effects) of potential drugs that inhibit kinases or phosphatases. For each deletion strain, we assessed changes in growth speed (*41*) and morphological features (table S8) (*42*). Despite 97 of the deletion strains showing reproducible responses in the phosphorylation network, only 9 mutants showed a strong effect on growth speed, and the total was 23 if strong changes in morphological features were also included (Fig. 3A). Conversely, 11 of the 27 kinases and phosphatases that had an undetectable,

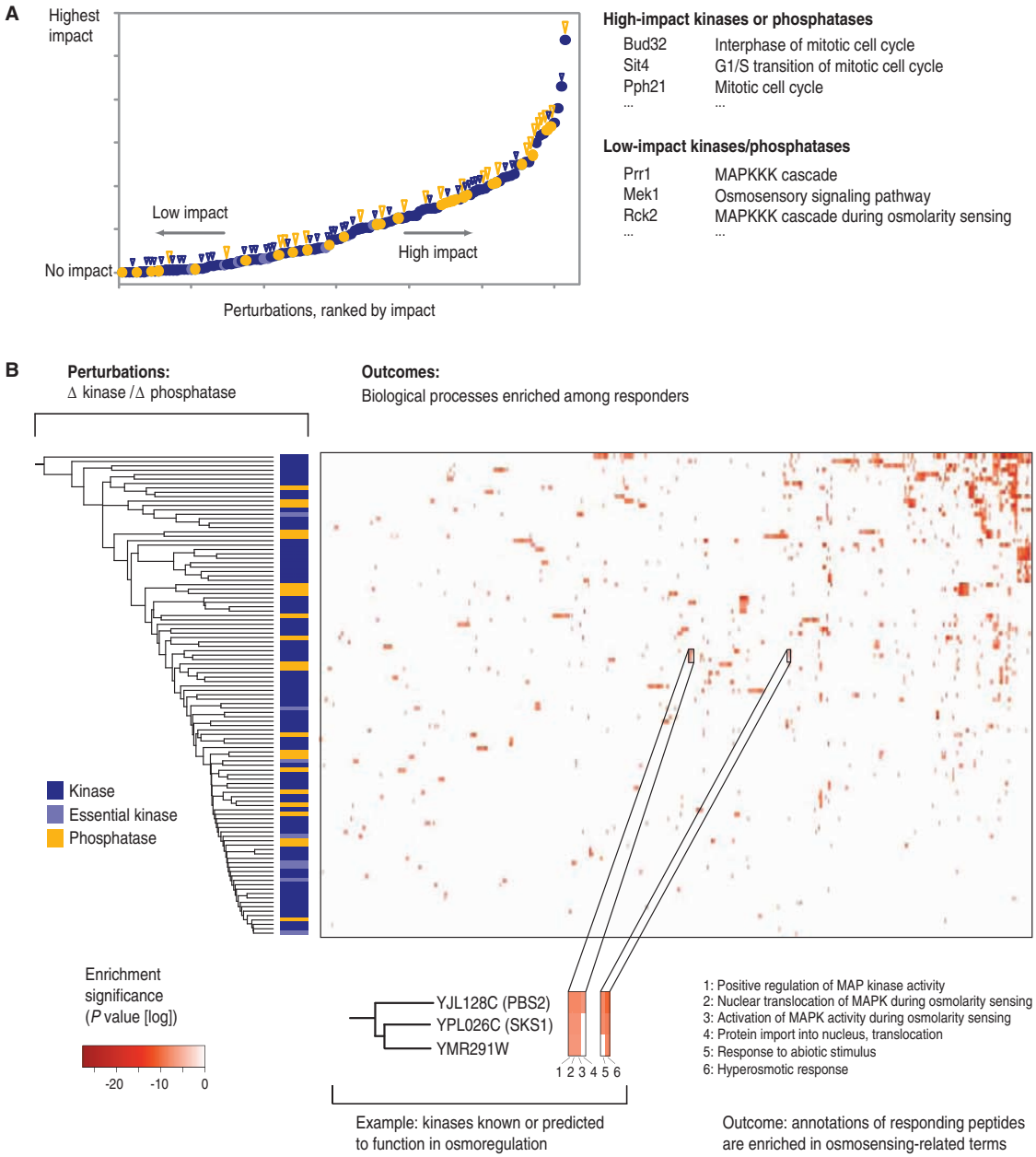


Fig. 3. (A) Phosphoproteome-wide impact of each kinase and phosphatase. For all kinases and phosphatases, we computed the fraction of phosphopeptides affected relative to the total number of phosphopeptides affected by all kinases and phosphatases. The kinases and phosphatases were then ranked accordingly. Blue circles represent kinases, light blue circles represent essential kinases, and golden circles represent phosphatases. A large golden triangle indicates a strong growth or morphological phenotype of a given mutant, whereas a small blue triangle represents a weak growth or morphological phenotype of a given mutant. Right side: examples of kinases that showed either a low or a high effect on the phosphoproteome regions, together with their known cellular functions. (B) For each kinase and phosphatase, the biological processes enriched among their regulated phosphoproteins were computed. Each column corresponds to a biological process, whereas each row corresponds to a given kinase or phosphatase (kinases are depicted in blue, essential kinases in light blue, and phosphatases in gold). The color scale denotes the statistical significance of the observed enrichment. Magnified inset: an example for three clustered kinases, for which a related set of processes is observed enriched among their substrates.

or only minimal, effect on the section of the phosphoproteome measured in this study showed a phenotype, among them, the kinase Elm1 (*43*), which showed a strong morphological phenotype. However, many strong morphological phenotypes were indeed observed in mutants that showed a strong change in the phosphoproteome, but the results were nevertheless surprising because they indicated that strong phenotypes were not necessarily reflected in the status of the phosphoproteome, as exemplified by Elm1 and other enzymes. Perhaps, in some cases, compensatory effects (visible at the level of the phosphoproteome) were precisely what prevented the occurrence of strong phenotypic consequences, as exemplified by the lack of correlation between the growth phenotypes and the changes in the phosphoproteome. This observation is particularly relevant because, first, cancer cells might display in some regards increased compensatory power, and second, kinase inhibitors that are specific for a target in vivo might not necessarily result in a cellular phenotype.

DISCUSSION

Our study delineates the responses of the system-wide cellular phosphorylation network upon systematic inactivation of individual kinases or phosphatases. Because the phosphorylation network is one of the main cellular backbones for the processing of information and the implementation of cellular responses, it is highly dynamic. Our measured behavior is only a single snapshot of a large number of possible outcomes, which were constrained by the growth and experimental conditions that we chose.

The first surprising observation that we made was that 7550 phosphopeptides were consistently identified but did not show a substantial amount of regulation. This may be due to, first, our cutoffs being conservative; thus, many putative regulatory events may not have been reproducible or strong enough to be deemed substantial. Second, 22% of the kinase and phosphatase mutants could not be analyzed, mainly because the corresponding genes are essential for cellular viability. Perhaps their essentiality is at least partly due to a generally higher impact on the phosphoproteome, as indicated recently (*10*), or because their substrates need to be phosphorylated constitutively. Third, in yeast, a large number of paralogous kinase isoforms exist (for example, Tpk1, Tpk2, and Tpk3). Given this, it is reasonable to expect some overlap or redundancy in substrates, which could lead to a considerable number of phosphorylation sites that would appear unregulated as long as only one of the paralogous duplicates was deleted. Fourth, the yeast populations that we analyzed consisted in a strict sense of many mixed subpopulations (for example, cells in different cell cycle states), and it can be assumed that an identical phosphorylation site can become phosphorylated by different kinases during the cell cycle. Therefore, analyzing deletions of single kinases or phosphatases would only manifest in slight, if any, regulation for such sites; for example, a cell cycle phase–specific regulation is masked by all cells that are not in that particular phase at any given time point. Fifth, we also analyzed whether the regulated and nonregulated phosphopeptides fell into different protein abundance classes (for example, the nonregulated are of low abundance and therefore regulation is more difficult to observe), but this was not the case. Overall, it is likely that all five possible explanations contribute to the observed result.

Another finding of this study was the unexpectedly strong dominance of indirect effects (as opposed to direct molecular target effects), which were often without a resulting strong cellular phenotype. To some extent, this observation fits with a view of signaling networks having to be highly flexible and redundant to respond to an ever-changing environment while maintaining stable cellular states (*44*). This constrains the architecture of the system, as described by the "law of requisite variety" (*45*, *46*), a fundamental law in systems control theory. It states that stable systems have to encode a number of control states that is higher than or equal to the number of states to be controlled. Considering that for each cell the space of "environmental states" is enormous, consequently, also the cellular "control variable space" must have an equal or greater size. The combinatorial possibilities of the phosphoproteome seem to ideally fulfill this demand (*44*).

An alternative explanation for this observation might also be found in the theory of Neutral Evolution (*47*). It is possible that only a small number of the observed phosphorylation events are actually relevant for the function and survival of the cell, whereas most phosphorylation events would simply have no effect, or at least have no negative effect, on the cell. As a result, such phosphorylation sites would not be counterselected during evolution. The data generated in this study do not, by themselves, support or refute this hypothesis. Finally, the low correlation between phenotype and the degree of change in the phosphoproteome may have been affected by the growth conditions chosen here, the lack of sensitivity of the phenotypic assays, or the possibility that the phosphoproteomics data were not sampled deeply enough to find such correlations.

In addition to revealing insights into the architecture of cellular signaling, our data set also describes the proteome-wide functional states of yeast cells; this might be useful for determining diagnostic markers for stress conditions, functional states of key pathways, or the activity of a given kinase or phosphatase. These markers could be used in conjunction with targeted proteomics approaches to not only study basic biological processes but also determine how a given pharmacological intervention would affect the cellular signaling network.

With targeted proteomics methods, not only can the cellular information flux under many conditions be observed, at high throughput, but this approach also enables us to understand for all phosphorylation sites whether the observed change is a "true" regulation event or simply as a result of a change in protein abundance (*48*–*50*) because both the phosphopeptide and several proteotypic peptides corresponding to the protein could be relatively or absolutely quantified, thus determining the phosphorylation site occupancy and regulation. Overall, our data provide global starting points, and constraints, toward understanding the complexity of phosphorylation regulation in yeast and other organisms. In the future, the results should be complemented by similar data for specific cellular conditions, time courses, or small-molecule interventions, thereby sharpening—step by step—our view of the events in the phosphorylation network. The ensuing insights in general design rules and motifs in cellular information processing will be essential for our ability to develop kinase-based drugs in an informed way.

MATERIALS AND METHODS

The generated LC-MS/MS phosphoproteome maps (table S2), an overview of the generated data (table S12), and the statistical methods used for their analysis are explained in detail in the Supplementary Materials. We have made available all kinase/phosphatase-responder relations in a user-friendly way in the recently described PhosphoPep database (*30*, *51*) (http://www.phosphopep.org). All yeast strains used here can be supplied upon request in a 96-well plate format (table S13).

SUPPLEMENTARY MATERIALS

www.sciencesignaling.org/cgi/content/full/3/153/rs4/DC1
Materials and Methods
Fig. S1. Power of the analysis approach.
Fig. S2. Topological properties of the protein phosphorylation network.
Fig. S3. Abundance distribution of responder phosphoproteins (proteins that contain "regulated" phosphopeptides).
Fig. S4. Ratio of phosphopeptides that are reduced or increased in abundance.
Fig. S5. Regulation of phosphopeptides versus regulation of protein abundance.
Fig. S6. Regulation of phosphopeptides versus regulation of protein abundance.
Fig. S7. Regulation of phosphopeptides that map to the same protein.

Table S1. List of enzymes.
Table S2. False discovery rate of peptide identification and specificity of phosphopeptide enrichment for each analyzed phosphorylation pattern.
Table S3. Information on phosphopeptides and phosphoproteins.
Table S4. Significant coregulation of kinases and phosphatases.
Table S5. Overlap of data from this study with other data sets.
Table S6. Confirmed STRING interactions.
Table S7. Overlap of possible direct targets with other data sets.
Table S8. Effects of each kinase and phosphatase on the phosphoproteome.
Table S9. Enrichment of biological processes among the low-impact kinases (bottom half).
Table S10. Enrichment of biological processes among the high-impact kinases (top half).
Table S11. GO terms.
Table S12. Overview of the entire data set.
Table S13. Information on yeast strains.
References

REFERENCES AND NOTES

1. P. Cohen, Protein kinases—the major drug targets of the twenty-first century? *Nat. Rev. Drug Discov.* **1**, 309–315 (2002).
2. T. Hunter, Signaling—2000 and beyond. *Cell* **100**, 113–127 (2000).
3. J. M. Irish, R. Hovland, P. O. Krutzik, O. D. Perez, Ø. Bruserud, B. T. Gjertsen, G. P. Nolan, Single cell profiling of potentiated phospho-protein networks in cancer cells. *Cell* **118**, 217–228 (2004).
4. L. Pelkmans, E. Fava, H. Grabner, M. Hannus, B. Habermann, E. Krausz, M. Zerial, Genome-wide analysis of human kinases in clathrin- and caveolae/raft-mediated endocytosis. *Nature* **436**, 78–86 (2005).
5. D. G. Savage, K. H. Antman, Imatinib mesylate—a new oral targeted therapy. *N. Engl. J. Med.* **346**, 683–693 (2002).
6. J. Ptacek, G. Devgan, G. Michaud, H. Zhu, X. Zhu, J. Fasolo, H. Guo, G. Jona, A. Breitkreutz, R. Sopko, R. R. McCartney, M. C. Schmidt, N. Rachidi, S. J. Lee, A. S. Mah, L. Meng, M. J. Stark, D. F. Stern, C. De Virgilio, M. Tyers, B. Andrews, M. Gerstein, B. Schweitzer, P. F. Predki, M. Snyder, Global analysis of protein phosphorylation in yeast. *Nature* **438**, 679–684 (2005).
7. R. Linding, L. J. Jensen, G. J. Ostheimer, M. A. van Vugt, C. Jorgensen, I. M. Miron, F. Diella, K. Colwill, L. Taylor, K. Elder, P. Metalnikov, V. Nguyen, A. Pasculescu, J. Jin, J. G. Park, L. D. Samson, J. R. Woodgett, R. B. Russell, P. Bork, M. B. Yaffe, T. Pawson, Systematic discovery of in vivo phosphorylation networks. *Cell* **129**, 1415–1426 (2007).
8. D. Fiedler, H. Braberg, M. Mehta, G. Chechik, G. Cagney, P. Mukherjee, A. C. Silva, M. Shales, S. R. Collins, S. van Wageningen, P. Kemmeren, F. C. Holstege, J. S. Weissman, M. C. Keogh, D. Koller, K. M. Shokat, N. J. Krogan, Functional organization of the *S. cerevisiae* phosphorylation network. *Cell* **136**, 952–963 (2009).
9. G. M. Rishton, Failure and success in modern drug discovery: Guiding principles in the establishment of high probability of success drug discovery organizations. *Med. Chem.* **1**, 519–527 (2005).
10. L. J. Holt, B. B. Tuch, J. Villen, A. D. Johnson, S. P. Gygi, D. O. Morgan, Global analysis of Cdk1 substrate phosphorylation sites provides insights into evolution. *Science* **325**, 1682–1686 (2009).
11. A. Huber, B. Bodenmiller, A. Uotila, M. Stahl, S. Wanka, B. Gerrits, R. Aebersold, R. Loewith, Characterization of the rapamycin-sensitive phosphoproteome reveals that Sch9 is a central coordinator of protein synthesis. *Genes Dev.* **23**, 1929–1943 (2009).
12. J. V. Olsen, B. Blagoev, F. Gnad, B. Macek, C. Kumar, P. Mortensen, M. Mann, Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* **127**, 635–648 (2006).
13. E. S. Lander, L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. Lehoczky, R. LeVine, P. McEwan, K. McKernan, J. Meldrim, J. P. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, N. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J. C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, R. H. Waterston, R. K. Wilson, L. W. Hillier, J. D. McPherson, M. A. Marra, E. R. Mardis, L. A. Fulton, A. T. Chinwalla, K. H. Pepin, W. R. Gish, S. L. Chissoe, M. C. Wendl, K. D. Delehaunty, T. L. Miner, A. Delehaunty, J. B. Kramer, L. L. Cook, R. S. Fulton, D. L. Johnson, P. J. Minx, S. W. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J. F. Cheng, A. Olsen, S. Lucas, C. Elkin, E. Uberbacher, M. Frazier, R. A. Gibbs, D. M. Muzny, S. E. Scherer, J. B. Bouck, E. J. Sodergren, K. C. Worley, C. M. Rives, J. H. Gorrell, M. L. Metzker, S. L. Naylor, R. S. Kucherlapati, D. L. Nelson, G. M. Weinstock, Y. Sakaki, A. Fujiyama, M. Hattori, T. Yada, A. Toyoda, T. Itoh, C. Kawagoe, H. Watanabe, Y. Totoki, T. Taylor, J. Weissenbach, R. Heilig, W. Saurin, F. Artiguenave, P. Brottier, T. Bruls, E. Pelletier, C. Robert, P. Wincker, D. R. Smith, L. Doucette-Stamm, M. Rubenfield, K. Weinstock, H. M. Lee, J. Dubois, A. Rosenthal, M. Platzer, G. Nyakatura, S. Taudien, A. Rump, H. Yang, J. Yu, J. Wang, G. Huang, J. Gu, L. Hood, L. Rowen, A. Madan, S. Qin, R. W. Davis, N. A. Federspiel, A. P. Abola, M. J. Proctor, R. M. Myers, J. Schmutz, M. Dickson, J. Grimwood, D. R. Cox, M. V. Olson, R. Kaul, C. Raymond, N. Shimizu, K. Kawasaki, S. Minoshima, G. A. Evans, M. Athanasiou, R. Schultz, B. A. Roe, F. Chen, H. Pan, J. Ramser, H. Lehrach, R. Reinhardt, W. R. McCombie, M. de la Bastide, N. Dedhia, H. Blocker, K. Hornischer, G. Nordsiek, R. Agarwala, L. Aravind, J. A. Bailey, A. Bateman, S. Batzoglou, E. Birney, P. Bork, D. G. Brown, C. B. Burge, L. Cerutti, H. C. Chen, D. Church, M. Clamp, R. R. Copley, T. Doerks, S. R. Eddy, E. E. Eichler, T. S. Furey, J. Galagan, J. G. Gilbert, C. Harmon, Y. Hayashizaki, D. Haussler, H. Hermjakob, K. Hokamp, W. Jang, L. S. Johnson, T. A. Jones, S. Kasif, A. Kaspryzk, S. Kennedy, W. J. Kent, P. Kitts, E. V. Koonin, I. Korf, D. Kulp, D. Lancet, T. M. Lowe, A. McLysaght, T. Mikkelsen, J. V. Moran, N. Mulder, V. J. Pollara, C. P. Ponting, G. Schuler, J. Schultz, G. Slater, A. F. Smit, E. Stupka, J. Szustakowski, D. Thierry-Mieg, J. Thierry-Mieg, L. Wagner, J. Wallis, R. Wheeler, A. Williams, Y. I. Wolf, K. H. Wolfe, S. P. Yang, R. F. Yeh, F. Collins, M. S. Guyer, J. Peterson, A. Felsenfeld, K. A. Wetterstrand, A. Patrinos, M. J. Morgan, P. de Jong, J. J. Catanese, K. Osoegawa, H. Shizuya, S. Choi, Y. J. Chen, International Human Genome Sequencing Consortium, Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
14. G. Manning, D. B. Whyte, R. Martinez, T. Hunter, S. Sudarsanam, The protein kinase complement of the human genome. *Science* **298**, 1912–1934 (2002).
15. W. H. Mager, J. Winderickx, Yeast as a model for medical and medicinal research. *Trends Pharmacol. Sci.* **26**, 265–273 (2005).
16. S. Lindquist, S. Krobitsch, L. Li, N. Sondheimer, Investigating protein conformation-based inheritance and disease in yeast. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **356**, 169–176 (2001).
17. T. F. Outeiro, P. J. Muchowski, Molecular genetics approaches in yeast to study amyloid diseases. *J. Mol. Neurosci.* **23**, 49–60 (2004).
18. T. Hunter, G. D. Plowman, The protein kinases of budding yeast: Six score and more. *Trends Biochem. Sci.* **22**, 18–22 (1997).
19. C. Widmann, S. Gibson, M. B. Jarpe, G. L. Johnson, Mitogen-activated protein kinase: Conservation of a three-kinase module from yeast to human. *Physiol. Rev.* **79**, 143–180 (1999).
20. J. A. Simon, A. Bedalov, Yeast as a model system for anticancer drug discovery. *Nat. Rev. Cancer* **4**, 481–492 (2004).
21. A. C. Bishop, J. A. Ubersax, D. T. Petsch, D. P. Matheos, N. S. Gray, J. Blethrow, E. Shimizu, J. Z. Tsien, P. G. Schultz, M. D. Rose, J. L. Wood, D. O. Morgan, K. M. Shokat, A chemical switch for inhibitor-sensitive alleles of any protein kinase. *Nature* **407**, 395–401 (2000).
22. B. Bodenmiller, L. N. Mueller, M. Mueller, B. Domon, R. Aebersold, Reproducible isolation of distinct, overlapping segments of the phosphoproteome. *Nat. Methods* **4**, 231–237 (2007).
23. M. R. Larsen, T. E. Thingholm, O. N. Jensen, P. Roepstorff, T. J. Jorgensen, Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol. Cell. Proteomics* **4**, 873–886 (2005).
24. L. N. Mueller, O. Rinner, A. Schmidt, S. Letarte, B. Bodenmiller, M. Y. Brusniak, O. Vitek, R. Aebersold, M. Muller, SuperHirn—a novel tool for high resolution LC-MS-based peptide/protein profiling. *Proteomics* **7**, 3470–3480 (2007).
25. M. Y. Brusniak, B. Bodenmiller, D. Campbell, K. Cooke, J. Eddes, A. Garbutt, H. Lau, S. Letarte, L. N. Mueller, V. Sharma, O. Vitek, N. Zhang, R. Aebersold, J. D. Watts, Corra: Computational framework and tools for LC-MS discovery and targeted mass spectrometry-based proteomics. *BMC Bioinformatics* **9**, 542 (2008).
26. A. Keller, A. I. Nesvizhskii, E. Kolker, R. Aebersold, Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* **74**, 5383–5392 (2002).
27. J. E. Elias, S. P. Gygi, Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat. Methods* **4**, 207–214 (2007).
28. G. Smyth, Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* **3**, Article 3 (2004).
29. Y. Benjamini, Y. Hochberg, Controlling the false discovery rate—a practical and powerful approach to multiple testing. *J. R. Statist. Soc. Series B Stat. Methodol.* **57**, 289–300 (1995).
30. B. Bodenmiller, J. Malmstrom, B. Gerrits, D. Campbell, H. Lam, A. Schmidt, O. Rinner, L. N. Mueller, P. T. Shannon, P. G. Pedrioli, C. Panse, H. K. Lee, R. Schlapbach, R. Aebersold, PhosphoPep—a phosphoproteome resource for systems biology research in *Drosophila* Kc167 cells. *Mol. Syst. Biol.* **3**, 139 (2007).
31. D. Schwartz, M. F. Chou, G. M. Church, Predicting protein post-translational modifications using meta-analysis of proteome scale data sets. *Mol. Cell. Proteomics* **8**, 365–379 (2009).
32. B. J. Breitkreutz, C. Stark, T. Reguly, L. Boucher, A. Breitkreutz, M. Livstone, R. Oughtred, D. H. Lackner, J. Bahler, V. Wood, K. Dolinski, M. Tyers, The BioGRID Interaction Database: 2008 update. *Nucleic Acids Res.* **36**, D637–D640 (2008).
33. L. J. Jensen, M. Kuhn, M. Stark, S. Chaffron, C. Creevey, J. Muller, T. Doerks, P. Julien, A. Roth, M. Simonovic, P. Bork, C. von Mering, STRING 8—a global view on proteins

and their functional interactions in 630 organisms. *Nucleic Acids Res.* **37**, D412–D416 (2009).
34. J. Mok, P. M. Kim, H. Y. Lam, S. Piccirillo, X. Zhou, G. R. Jeschke, D. L. Sheridan, S. A. Parker, V. Desai, M. Jwa, E. Cameroni, H. Niu, M. Good, A. Remenyi, J. L. Ma, Y. J. Sheu, H. E. Sassi, R. Sopko, C. S. Chan, C. De Virgilio, N. M. Hollingsworth, W. A. Lim, D. F. Stern, B. Stillman, B. J. Andrews, M. B. Gerstein, M. Snyder, B. E. Turk, Deciphering protein kinase specificity through large-scale analysis of yeast phosphorylation site motifs. *Sci. Signal.* **3**, ra12 (2010).
35. C. Stark, T. C. Su, A. Breitkreutz, P. Lourenco, M. Dahabieh, B. J. Breitkreutz, M. Tyers, I. Sadowski, PhosphoGRID: A database of experimentally verified in vivo protein phosphorylation sites from the budding yeast *Saccharomyces cerevisiae*. *Database* (*Oxford*) **2010**, bap026 (2010).
36. J. A. Ubersax, E. L. Woodbury, P. N. Quang, M. Paraz, J. D. Blethrow, K. Shah, K. M. Shokat, D. O. Morgan, Targets of the cyclin-dependent kinase Cdk1. *Nature* **425**, 859–864 (2003).
37. Y. Hirata, T. Andoh, T. Asahara, A. Kikuchi, Yeast glycogen synthase kinase-3 activates Msn2p-dependent transcription of stress responsive genes. *Mol. Biol. Cell* **14**, 302–312 (2003).
38. A. G. Hinnebusch, K. Natarajan, Gcn4p, a master regulator of gene expression, is controlled at multiple levels by diverse signals of starvation and stress. *Eukaryot. Cell* **1**, 22–32 (2002).
39. D. E. Sterner, J. M. Lee, S. E. Hardin, A. L. Greenleaf, The yeast carboxyl-terminal repeat domain kinase CTDK-I is a divergent cyclin–cyclin-dependent kinase complex. *Mol. Cell. Biol.* **15**, 5716–5724 (1995).
40. J. H. Grose, T. L. Smith, H. Sabic, J. Rutter, Yeast PAS kinase coordinates glucose partitioning in response to metabolic and cell integrity signaling. *EMBO J.* **26**, 4824–4830 (2007).
41. M. E. Hillenmeyer, E. Fung, J. Wildenhain, S. E. Pierce, S. Hoon, W. Lee, M. Proctor, R. P. St Onge, M. Tyers, D. Koller, R. B. Altman, R. W. Davis, C. Nislow, G. Giaever, The chemical genomic portrait of yeast: Uncovering a phenotype for all genes. *Science* **320**, 362–365 (2008).
42. Y. Ohya, J. Sese, M. Yukawa, F. Sano, Y. Nakatani, T. L. Saito, A. Saka, T. Fukuda, S. Ishihara, S. Oka, G. Suzuki, M. Watanabe, A. Hirata, M. Ohtani, H. Sawai, N. Fraysse, J. P. Latgé, J. M. François, M. Aebi, S. Tanaka, S. Muramatsu, H. Araki, K. Sonoike, S. Nogami, S. Morishita, High-dimensional and large-scale phenotyping of yeast mutants. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 19015–19020 (2005).
43. M. J. Blacketer, C. M. Koehler, S. G. Coats, A. M. Myers, P. Madaule, Regulation of dimorphism in *Saccharomyces cerevisiae*: Involvement of the novel protein kinase homolog Elm1p and protein phosphatase 2A. *Mol. Cell. Biol.* **13**, 5567–5581 (1993).
44. M. Thomson, J. Gunawardena, Unlimited multistability in multisite phosphorylation systems. *Nature* **460**, 274–277 (2009).
45. W. R. Ashby, *An Introduction to Cybernetics* (Chapman & Hall Ltd., London, 1956).
46. W. R. Ashby, Requisite variety and its implications for the control of complex systems. *Cybernetica* **1**, 83–99 (1958).
47. M. Kimura, Evolutionary rate at the molecular level. *Nature* **217**, 624–626 (1968).
48. V. Lange, P. Picotti, B. Domon, R. Aebersold, Selected reaction monitoring for quantitative proteomics: A tutorial. *Mol. Syst. Biol.* **4**, 222 (2008).
49. P. Picotti, R. Aebersold, B. Domon, The implications of proteolytic background for shotgun proteomics. *Mol. Cell. Proteomics* **6**, 1589–1598 (2007).
50. P. Picotti, H. Lam, D. Campbell, E. W. Deutsch, H. Mirzaei, J. Ranish, B. Domon, R. Aebersold, A database of mass spectrometric assays for the yeast proteome. *Nat. Methods* **5**, 913–914 (2008).
51. B. Bodenmiller, D. Campbell, B. Gerrits, H. Lam, M. Jovanovic, P. Picotti, R. Schlapbach, R. Aebersold, PhosphoPep—a database of protein phosphorylation sites in model organisms. *Nat. Biotechnol.* **26**, 1339–1340 (2008).
52. **Acknowledgments:** We thank the whole team of the Functional Genomics Center Zurich (FGCZ) for fruitful discussions. We thank C. Zheng, Department of Statistics, Purdue University, for help with use and interpretation of the Limma package. **Funding:** This project was funded in part by ETH Zurich; Federal funds from the National Heart, Lung, and Blood Institute, NIH, under contract no. N01-HV-28179; the PhosphoNetX project of SystemsX.ch, the Swiss initiative for systems biology; and the European Research Council (grant ERC-2008-AdG 233226) to R.A. Work at the FGCZ and at the von Mering laboratory has been supported by the University Research Priority Program Systems Biology and Functional Genomics of the University of Zurich. B.G. is supported by the Bonizzi-Theler Foundation. C.K. is supported by a Marie-Heim Vögtlin fellowship from the Swiss National Science Foundation (SNF). B.B. was the recipient of fellowships by the Boehringer Ingelheim Fonds, the SNF, and the EMBO Long-Term Fellowship. **Author contributions:** B.B. coordinated the project, conducted most of the experimental work, and contributed to data analysis and writing the manuscript; S.W. and C.v.M. conducted downstream bioinformatics analysis and contributed to writing the manuscript; C.K. and M.P. generated the kinase and phosphatase deletion mutants; J.U. and R.L. generated and characterized a number of the analog-sensitive kinases; D.C. implemented all data into the PhosphoPep database; P.G.P., G.P.N., and A.I.N. contributed to bioinformatics analysis of data; P.P. supported generation of yeast samples; B.G. and B.R. supported LC-MS/MS method optimization and data acquisition; R.S. consulted and coordinated work at the FGCZ; H.L. processed the collected MS/MS spectra of phosphopeptides, validated their identification, and created high-quality spectral libraries of yeast phosphopeptides for distribution; A.C.-L. determined the morphological phenotypes of the yeast cells; M.-Y.B. modified the Corra interface and processed feature-aligned and postprocessed data sets for differentially expressed feature detection and probability assignment; O.V. performed statistical analyses of the data; C.Z. and K.M.S. synthesized kinase inhibitors; and R.A. conceptualized the study and contributed to writing the manuscript. **Competing interests:** The authors declare that they have no competing interests.

Submitted 14 May 2010
Accepted 3 December 2010
Final Publication 21 December 2010
10.1126/scisignal.2001182
Citation: B. Bodenmiller, S. Wanka, C. Kraft, J. Urban, D. Campbell, P. G. Pedrioli, B. Gerrits, P. Picotti, H. Lam, O. Vitek, M.-Y. Brusniak, B. Roschitzki, C. Zhang, K. M. Shokat, R. Schlapbach, A. Colman-Lerner, G. P. Nolan, A. I. Nesvizhskii, M. Peter, R. Loewith, C. von Mering, R. Aebersold, Phosphoproteomic analysis reveals interconnected system-wide responses to perturbations of kinases and phosphatases in yeast. *Sci. Signal.* **3**, rs4 (2010).

Post-processing of SuperHirn results to avoid signals caused by missed feature alignment

Occasionally, SuperHirn fails to correctly align some of the phosphopeptide features from multiple LC-MS/MS runs. If such an alignment failure occurs when comparing mutant and wild-type samples, then it can lead to erroneous inference of a "reliable" change in phosphopeptide abundance in response to a mutation (for example, in cases where the peptide feature happens to be aligned in all three wild-type samples, but not in the mutant samples). To avoid such false signals, we identified potential mis-aligned features in a post-processing step and, in each case, merged them to a single representative feature. Essentially, we identified mis-aligned features based on similarities in mass and retention time (allowing for a certain "window" of tolerance). In a first step, we only considered features that had MS2 information and a PeptideProphet score of >0.9. We created clusters of potentially mis-aligned features by grouping those with the same peptide sequence, mass, and retention time. For both mass and retention time, we permitted tolerances of 15 ppm for mass and 7 min for retention time; the latter tolerance was increased to 15 min at the beginning and end of each LC separation run. MS2 sequences had to be identical, except with respect to the exact position of the phosphorylated residue (allowing a shift of a maximum of 5 residues). In a second step, we extended the clusters by including non-MS2-features. For these, more stringent tolerance intervals were used: 3 ppm mass tolerance, and a 3- or 5-min retention time tolerance. Finally, the features of every cluster were merged to a single representative feature. The intensity values were added run-wise, the average mass of all features was used as the new aligned feature mass, and for retention time and charge, the full range was provided. The MS2 data of an aligned feature were represented by a non-redundant list of all of the encountered sequences. For features that were apparently partitioned exactly along the wild-type–mutant distinction (that is, observed in the wild-type, but never observed in the mutant), we intentionally increased the window of tolerance to make sure that these observations were real (that is, we erred to be on the safe side here, even if this occasionally meant the loss of a real signal). For these features, tolerances were 25 ppm mass tolerance and 15 or 20 min retention time in the first step, and 3 ppm mass tolerance and a 5- or 7-min retention time in the second step. All tolerance-values stated above were chosen based on manual inspection of the data, with the simple rule that distinct features with matching MS2 identification should be merged, but that features with incompatible MS2 identifications should not be merged. In the entire post-processing pipeline, only features with a signal-to-noise-ratio better than 10 were considered (for features that seemed to appear or disappear in the comparison between wild-type and mutant, the signal-to-noise-ratio threshold was increased to 60), assuming as background noise level in a given run the average of the 50 lowest signals as detected by SuperHirn. With respect to the MS2 sequences, half-tryptic peptides were discarded and all other peptides were merged with the corresponding peptide. Finally, we performed two manual validation steps before continuing with the data analyses. First, for very few cases, peptides indicating the gene product of a knock-out were observed. If an explanation for this observation could be found (low peptide probability or wrong peptide identification, sample carry over, mis-assignment, etc.) the entries were removed. If not, the master maps were omitted from further analyses. Second, we plotted the extracted ion chromatograms for all phosphopeptide ions that were only detected either in the wild-type or mutant replicates to manually validate their presence or absence ("vanished"/ "appeared").

Statistical significance of observed differential abundances of phosphopeptides

After post-processing of the SuperHirn output files, the phosphopeptides were separated into different statistical classes for further analysis. A class consisted either of phosphopeptides for which the MS1 signal was detected in all replicates (three times for the wild-type and three times for the mutant) or of phosphopeptides for which 1, 2, 3, or up to 5 signals were missing.

The category "3 signals missing" was further separated into phosphopeptides for which either (i) the signal was reproducibly present in either all wild-type or all mutant samples, or (ii) the signal was spread over all wild-type and mutant samples. Before statistical analysis, the missing data values were imputed with the integrated background noise as determined by the SuperHirn algorithm. These datasets were then further analyzed as described in detail previously (*15, 16*). In short, the employed software tool called Corra (*17*) wraps around the Limma (*16*) software package, which is available from the R-based project Bioconductor, and performs a statistical test comparing the abundances of a feature between the mutant and the wild type replicate samples. The test is an Empirical Bayes alternative to the Welsh *t* test, and is based on a moderated *t* statistic where the standard error is calculated with the information from all of the analyzed phosphopeptide features. Overall, the statistical analysis that we used assumes that the abundances of each feature follow a normal distribution across runs. (Normality analysis of our dataset showed that deviations to normality, even if present, were minor; the same was true for the assumption of equal distribution of our features.) Even if this assumption was not always fulfilled, the test statistics and their associated *P* values remain useful to detect changes in abundance. The test is conducted separately for each feature and each mutant.

When the number of replicate samples in the experiment is small, inference based on the *t* test may be somewhat unstable. The Empirical Bayes (or moderated) alternative to the *t* test, originally developed for gene expression microarrays (*16, 18*), is designed to remedy this problem. It combines the information regarding feature variability across all features, and improves both the sensitivity and specificity of finding the true changes in abundance across sample types. The resulting *P* value is then adjusted for multiple comparisons according to the procedure by Benjamini and Hochberg, which controls the FDR (*19*). A potential technical problem with the application of the procedure is the positive correlation in abundances of multiple peptides from the same differentially abundant protein, as well as an increased positive correlation between test statistics due to the use of the Empirical Bayes procedure. However, Benjamini *et al*. (*20*) demonstrated that calculations of the FDR are robust to positive inter-feature correlation. Therefore, the application of the procedure to our dataset was appropriate. As has been described for microarray data (*21*), a major factor for the reproducibility of a given regulation is the observed fold-change. Therefore, besides using the FDR cut-off of 0.015, we also required the fold-change in the abundance of a phosphopeptide to be equal or greater than a $\log_2$ of 1.5 to consider it as regulated. For features that were only seen in either the wild-type or the mutant ("vanishers" or "appearers"), this cut-off was raised to $\log_2 > 4$ to exclude the possibility that phosphopeptide ions just above (or below) the MS detection limit were detected to vanish (appear).

Changes in the extent of phosphorylation versus changes in protein abundance

For the first analysis (comparison of the changes in the abundance of phosphopeptides with those of non-phosphorylated peptides) the following deletion mutants were analyzed:

YCR079W
YGL059W
YGR040W
YHR135C
YKL171W
YOR231W
YBL088C

YDL079C
YDL101C
YGL180W
YGR092W
YJL128C
YLR248W
YNL307C
YPL140C
YPL141C

For the second analysis, all regulatory events were considered independently, provided that (i) at least one phosphopeptide was significantly regulated, and (ii) a second phosphopeptide mapped to the same phosphoprotein, irrespective of whether or not it was also observed as significantly regulated. Note that responses with opposite directionality are actually not expected to occur very frequently: They require both the disappearance and the appearance of specific phosphopeptides in response to deletion of the same kinase or phosphatase.

Global impact of kinases and phosphatases on the phosphoproteome

When computing the impact of a given kinase or phosphatase, we first normalized the number of regulated phosphopeptides to all of the identified phosphopeptides of that given kinase or phosphatase and to the total number of regulated phosphopeptides (see Fig. 2). The enzymes were then ranked according to their normalized impact. Note that the determined numbers were just an estimate, because of (i) experimental variations between analyzed batches and therefore differences in LC-MS maps; (ii) the incomplete covered phosphoproteome; and (iii) our computational approach of performing pair-wise comparisons between wild-type and mutant samples. We considered kinases or phosphatases as shown in Fig. 3A to be active if 10 or more phosphopeptides changed significantly in abundance upon deletion of the kinase or phosphatase or upon inhibition of the essential kinase.

Computation of the biological processes enriched in the inactive and active kinases and phosphatases

For this analysis, we only considered the gene deletion strains, because the inhibitable strains are not representative as a result of the variability in the impact on the phosphoproteome depending on the length of time of inhibition, for example. The enrichment analyses were performed with http://pipe.systemsbiology.net using the default settings (yeast proteome background).

Because the following terms are associated with the kinases and phosphatases themselves, we removed them from the results shown:

phosphate metabolic process
phosphorus metabolic process
protein amino acid phosphorylation
phosphorylation
post-translational protein modification
protein modification process
biopolymer modification
cellular protein metabolic process

protein metabolic process
cellular macromolecule metabolic process
signal transduction
cell communication
biological regulation
dephosphorylation
biopolymer metabolic process
protein amino acid dephosphorylation
regulation of biological process
regulation of cellular process
macromolecule metabolic process
regulation of biological process
regulation of cellular process
response to stimulus

Computation of the biological processes enriched among all responding proteins for a given kinase or phosphatase

For each kinase or phosphatase, the regulated phosphopeptides were mapped to their corresponding proteins. In case phosphopeptides mapped to several proteins, only those peptides mapping to homologous proteins (with equal biological processes) were retained. The enrichment of GO biological process terms was calculated with the hypergeometric test–based *GOstats* package version 2.8.0 and the yeast annotation package *org.Sc.sgd.db* version 2.2.6 from Bioconductor. Only GO terms with a *P* value <0.01 were considered. All values in table S11 are given in $\log_2$.

Determination of morphological phenotypes

The morphological phenotypes of all strains (table S4) were determined as described by Gordon *et al*. (*22*). Furthermore, phenotype data as measured by Ohya *et al*.(*23*) and growth speed as determined by Hillenmeyer *et al*. (*24*) were used for the analyses. A strain was considered to have a strong growth phenotype ("+") or very strong growth phenotype ("++") if it grew between to between 30 to 100%, >100% as fast/slow as the wild-type strain under the same growth conditions. A strain was considered to have a strong morphological phenotype ("+") or a very strong morphological phenotype ("++") if at least one or more than 10, respectively, of the 254 parameters had a significance equal to or smaller than 1×10^{-05}.

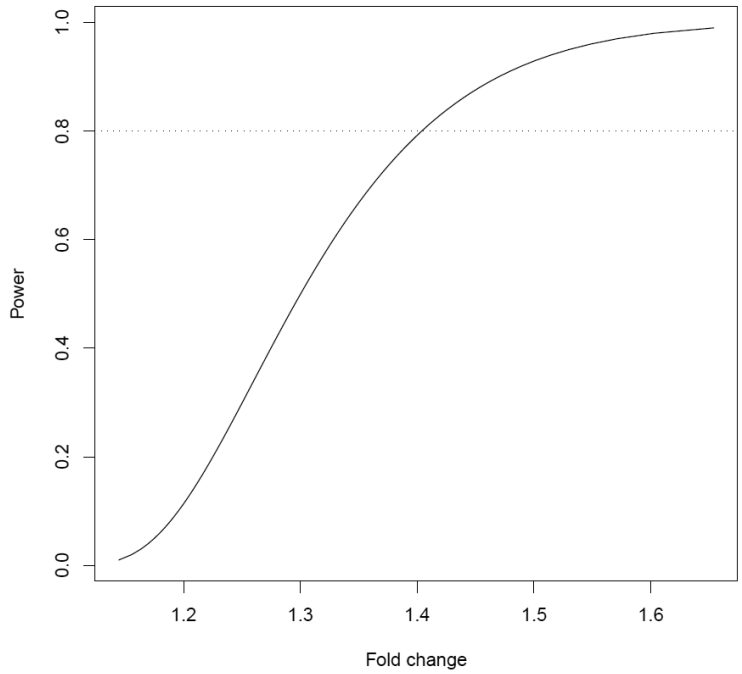


Fig. S1: Power of the analysis approach. The results of the Empirical Bayes approach were used to provide estimates of the power of our analysis, as described (*25*). Specifically, we took the median of posterior estimates of feature-specific variances provided by the Empirical Bayes *t* test, and the proportion of statistically significant changes in the datasets as input parameters. We set the FDR threshold to 0.1 and varied the expected fold-change between 1.0 and 1.7. The figure displays the power of the statistical test in this setting.

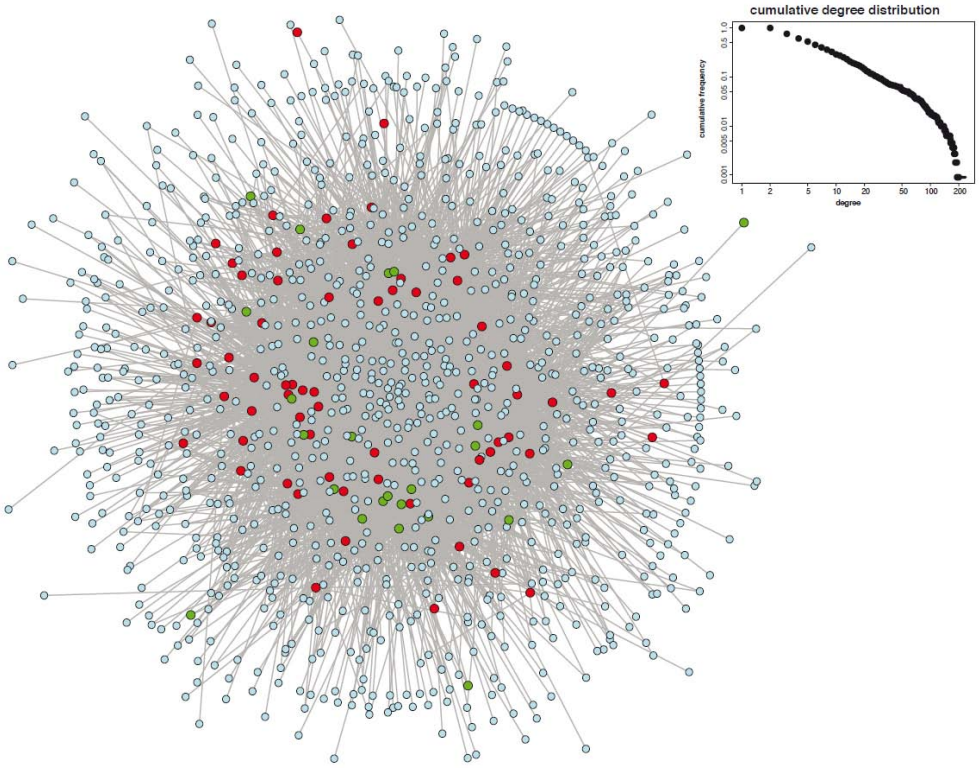


Fig. S2. Topological properties of the protein phosphorylation network. The network shown, which contains 1,088 nodes and 6,509 unique edges, is a summary view of the connectivity observed in our dataset. Kinases are shown in red, phosphatases in green, and responder proteins in light blue. Lines indicate substantial regulatory events. Inset: degree distribution of network connectivity. The x-axis shows the degree of connectivity of the network and the y-axis shows the cumulative frequency of the appearance of a given degree. The degree of a node is the number of edges connected to that node. The cumulative degree distribution was indicative of a scale-free topology. The average path length of the network is 3.1.

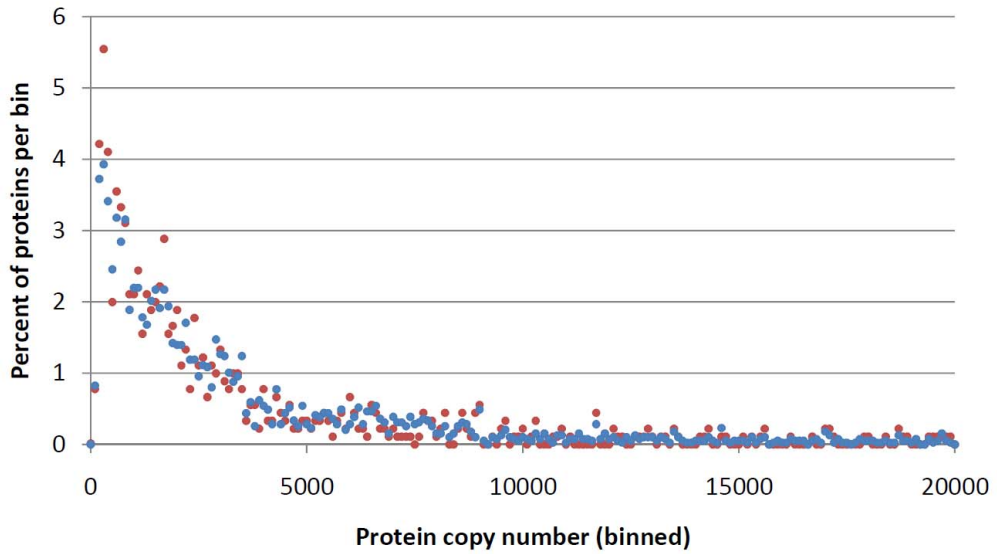


Fig. S3. Abundance distribution of responder phosphoproteins (proteins that contain "regulated" phosphopeptides). A comparison of the yeast proteome abundance distribution (blue) and the abundance of the regulated phosphoproteins as observed in this study (red, significance threshold FDR ≤ 0.015 and requiring at least a $\log_2$ 1.5-fold regulation or a full on or off response). The protein abundances were taken from Ghaemmaghami *et al.* (*26*). Proteins with more than 20,000 copies per cell are not displayed (the distribution of proteins with more than 20,000 copies per cell is similar between the analyzed phosphoprotein sets and the yeast proteome). The x-axis displays the protein copy number per cell, the y-axis the percentage of protein counts per copies per cell bin (with a bin size of 100) normalized by all of the proteins from the regulated phosphoprotein set or the data set of Ghaemmaghami *et al*. (*26*). The observations made for the responder phosphoproteins also held true if only the non-regulated phosphopeptides were analyzed, if the protein abundances were estimated by spectral counting as determined by Weiss *et al*. (*27*), or both.

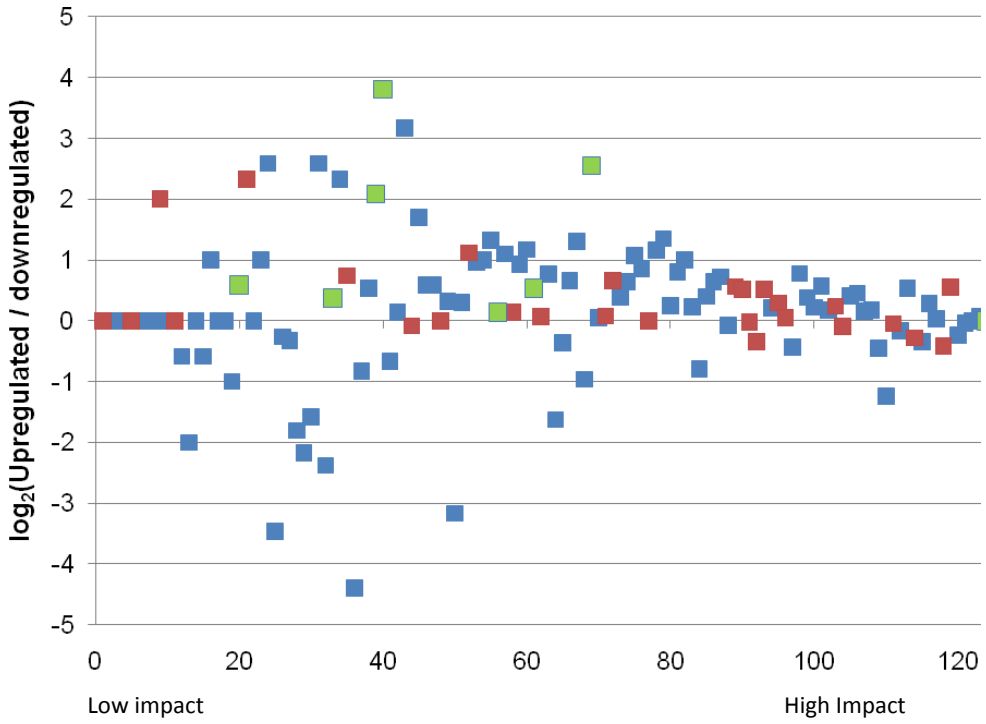


Fig. S4. Ratio of phosphopeptides that are reduced or increased in abundance. The y-axis shows the ratio [$\log_2$(number of increased phosphopeptides / number of decreased phosphopeptides)] and on the x-axis, the kinases and phosphatases are ordered according to the extent of their effect on the phosphoproteome (0 = lowest effect; 124 = highest effect, as calculated from the dataset in this study). The blue squares show the ratios for the kinases, the red squares show the ratios for the phosphatases, and the green squares show the ratios for the essential kinases. The higher ratios observed for the kinases and phosphatases with low activities were probably noise, because of the fewer regulated phosphopeptides observed. For this plot, relative regulation ("fold-changers") and complete disappearance ("on/off-responders", "vanishers") were not distinguished.

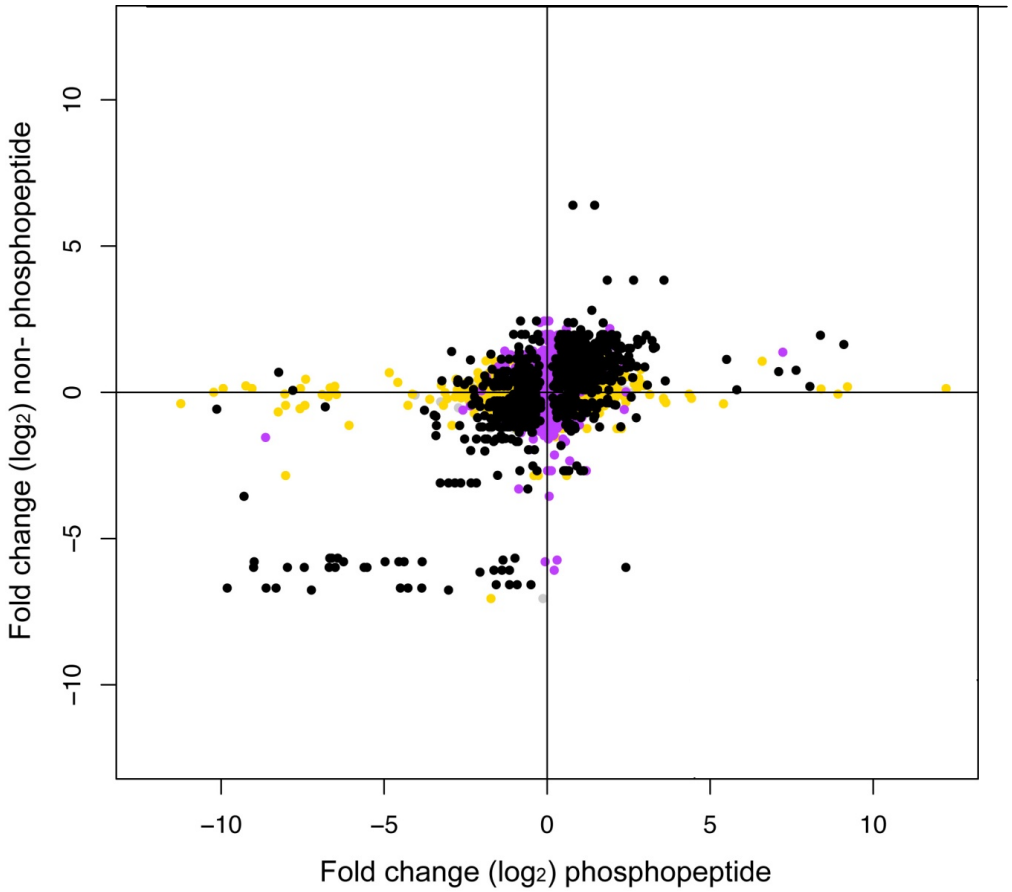


Fig. S5. Regulation of phosphopeptides versus regulation of protein abundance. Each dot corresponds to a measured phosphopeptide. The x-axis shows the fold-change (on a $\log_2$ scale) of the phosphopeptide and the y-axis shows the median change observed for all of the non-phosphorylated peptides that map to those particular proteins (both shown on a $\log_2$ scale). The color code illustrates whether the observed fold-change was significant in one or both of the measurements. ● denotes significant regulation detected in both phosphorylated and nonphosphorylated peptides; ● denotes significant regulation only for the phosphorylated peptide detected; ● denotes significant regulation only for the nonphosphorylated peptide detected; and ● denotes no significant regulation of any peptide detected.

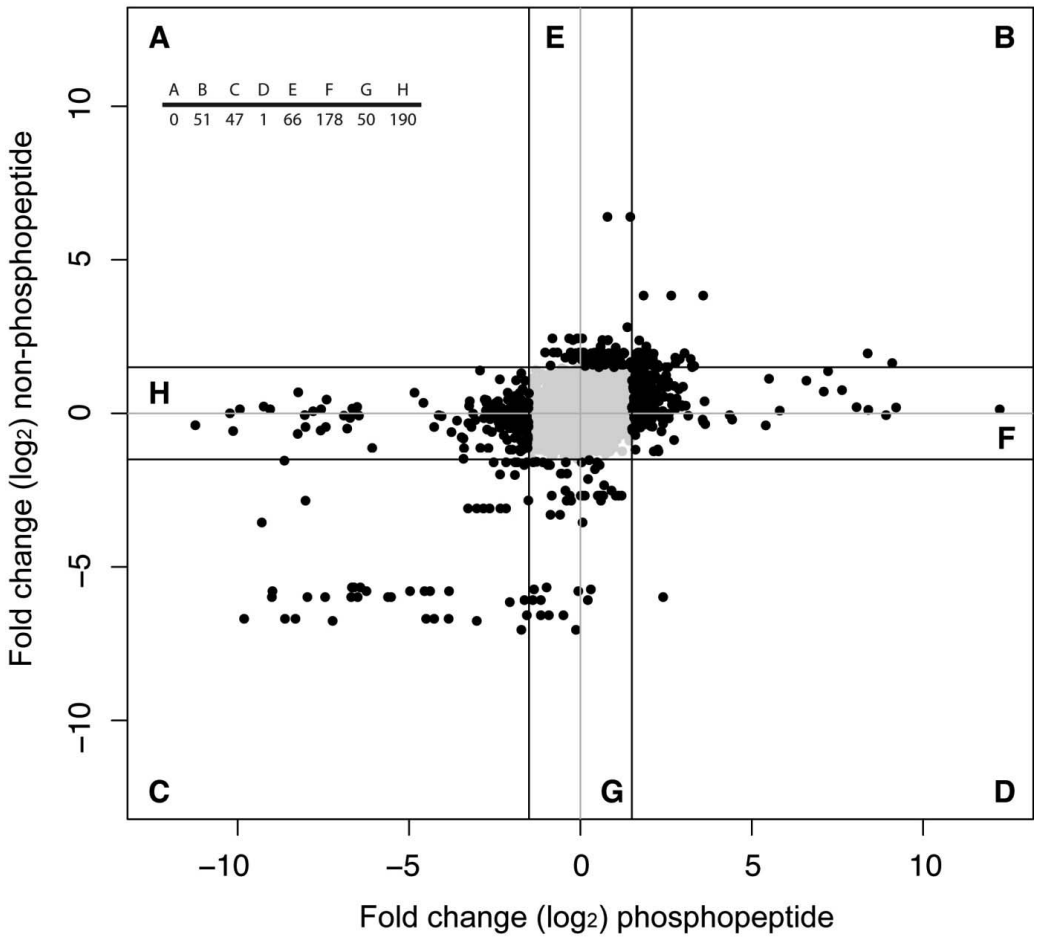


Fig. S6. Regulation of phosphopeptides versus regulation of protein abundance. The data shown are the same as those presented in fig. S5, but events that occurred below our standard fold-change cutoff are masked. This treatment enables us to see that the majority of the observed regulated changes in phosphopeptide abundance did not correspond to a change in protein abundance, but presumably in the occupancy of the phosphorylation sites. Note that this plot is based on a limited set of kinases (16 kinases), ranging from those that did not show any detectable impact on the phosphoproteome to those showing a large number of regulated phosphorylated peptides.

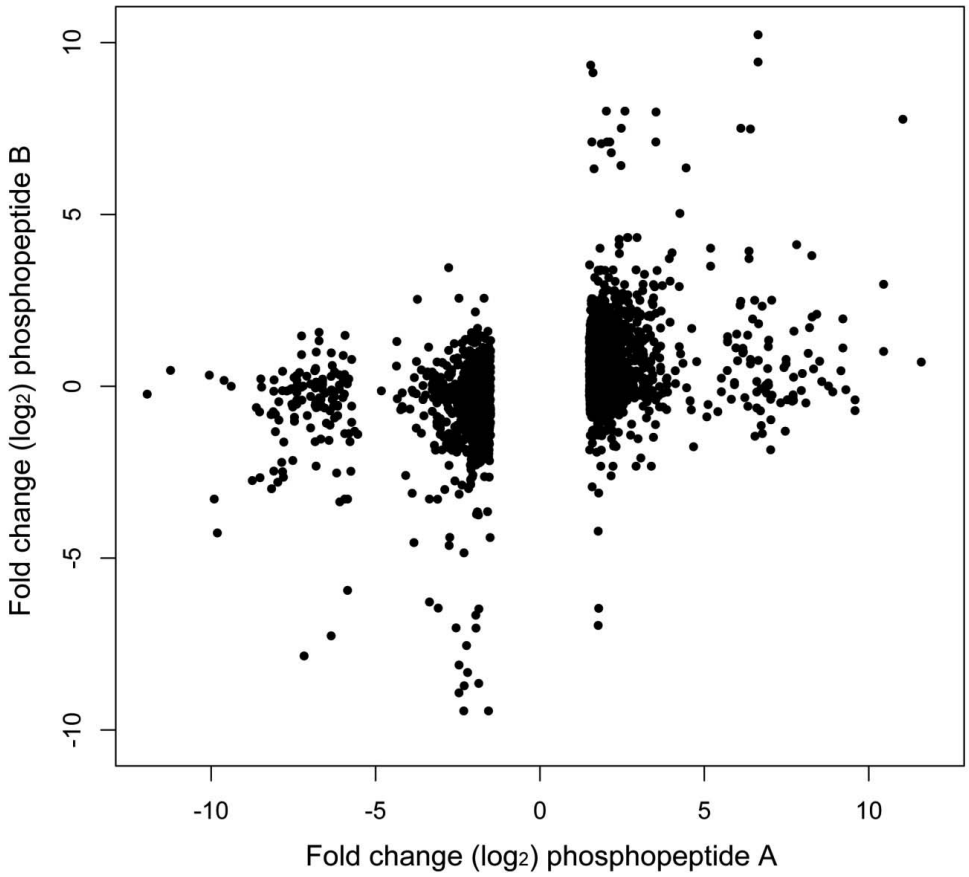


Fig. S7. Regulation of phosphopeptides that map to the same protein. On the x-axis are plotted phosphopeptides that were significantly regulated ($P < 0.015$; fold-change $\log_2 > 1.5$), whereas on the y-axis are plotted phosphopeptides that map to the corresponding phosphoprotein, irrespective of their significance or fold-change. In the majority of cases only one of the phosphopeptides that maps to a given phosphoprotein was regulated, indicating that the observed phosphorylation events were not due to a change in protein abundance.

Table S2. False discovery rate of peptide identification and specificity of phosphopeptide enrichment* for each analyzed phosphorylation pattern.

Kinase or phosphatase	Standard name	Group	FDR	% phosphopeptides*
YAL017W	PSK1	Other	0.05	86
YAR019C	CDC15	STE	0.03	79
YBL009W	ALK2		0.05	81
YBL016W	FUS3	CMGC	0.04	83
YBL056W	PTC3	STP	0.03	74
YBL088C	TEL1	Inositol Kinase	0.03	76
YBR028C		AGC	0.05	86
YBR059C	AKL1	Other	0.03	80
YBR097W	VPS15	Other	0.03	73
YBR125C	PTC4	STP	0.02	76
YBR160W	CDC28	CMGC	0.03	79
YBR274W	CHK1	Other	0.03	79
YBR276C	PPS1		0.04	83
YCR008W	SAT4	Other	0.04	82
YCR079W	PTC6	STP	0.04	85
YCR091W	KIN82	AGC	0.03	75
YDL006W	PTC1	STP	0.03	74
YDL025C	RTK1	Other	0.03	80
YDL028C	MPS1	Other	0.03	79
YDL047W	SIT4	STP	0.03	84
YDL079C	MRK1	CMGC	0.04	74
YDL101C	DUN1		0.03	74
YDL108W	KIN28	CMGC	0.03	78
YDL134C	PPH21	STP	0.05	77
YDL188C	PPH22	STP	0.05	85
YDL214C	PRR2	Other	0.04	85
YDL230W	PTP1	PTP	0.04	87
YDR075W	PPH3	STP	0.05	82
YDR122W	KIN1	CAMK	0.05	77
YDR247W	VHS1	Other	0.03	81
YDR283C	GCN2	Other	0.04	84
YDR466W	PKH3	AGC	0.03	83
YDR477W	SNF1	CAMK	0.03	87
YDR523C	SPS1	STE	0.03	81
YER129W	SAK1		0.04	82
YFR014C	CMK1	CAMK	0.05	86
YGL021W	ALK1		0.04	86
YGL059W	PKP2	Atypical PK	0.04	83

Table S2 continued.

Kinase or phosphatase	Standard name	Group	FDR	% phosphopeptides*
YGL179C	TOS3	CAMK	0.05	86
YGL180W	ATG1	Other	0.03	75
YGR040W	KSS1	CMGC	0.03	86
YGR123C	PPT1	STP	0.04	66
YGR188C	BUB1	Other	0.04	86
YGR203W	YCH1		0.04	82
YGR262C	BUD32	Microbial PK	0.03	76
YHR030C	SLT2	CMGC	0.06	82
YHR076W	PTC7		0.04	83
YHR082C	KSP1		0.05	78
YHR135C	YCK1	CKI	0.04	85
YIL035C	CKA1	Other	0.05	80
YIL042C	PKP1	Atypical PK	0.05	86
YIL095W	PRK1	Other	0.06	83
YIL113W	SDP1		0.04	75
YIR026C	YVH1		0.03	81
YJL095W	BCK1	STE	0.05	85
YJL106W	IME2	CMGC	0.02	75
YJL128C	PBS2	STE	0.02	71
YJL164C	TPK1	AGC	0.05	83
YJL165C	HAL5	Other	0.04	77
YJL187C	SWE1	Other	0.03	75
YJR059W	PTK2	Other	0.03	84
YJR066W	TOR1	Inositol Kinase	0.04	82
YKL048C	ELM1		0.03	73
YKL101W	HSL1	CAMK	0.04	82
YKL116C	PRR1	CAMK/EMK	0.05	83
YKL126W	YPK1	AGC	0.05	83
YKL139W	CTK1	CMGC	0.04	83
YKL161C	KDX1	CMGC	0.05	86
YKL166C	TPK3	AGC	0.04	84
YKL168C	KKQ8	Other	0.04	80
YKL171W	NNK1	Yeast PK	0.04	84
YKL198C	PTK1	Other	0.03	72
YLL010C	PSR1		0.05	77
YLL019C	KNS1	CMGC	0.05	86
YLR019W	PSR2		0.03	82
YLR096W	KIN2	CAMK	0.04	76

Table S2 continued.

Kinase or phosphatase	Standard name	Group	FDR	% phosphopeptides*
YLR113W	HOG1	CMGC	0.04	83
YLR240W	VPS34	Inositol Kinase	0.04	75
YLR248W	RCK2		0.03	74
YLR362W	STE11	STE	0.04	84
YLR433C	CNA1	STP	0.05	86
YML016C	PPZ1	STP	0.04	83
YML057W	CMP2	STP	0.05	86
YMR104C	YPK2	AGC	0.04	78
YMR139W	RIM11		0.03	82
YMR216C	SKY1	CMGC	0.04	83
YMR291W		CAMK	0.04	80
YNL020C	ARK1	Other	0.04	84
YNL032W	SIW14	PTP	0.03	86
YNL099C	OCA1		0.03	79
YNL154C	YCK2	CKI	0.04	65
YNL161W	CBK1	AGC	0.03	80
YNL298W	CLA4	STE	0.06	87
YNL307C	MCK1		0.04	73
YNR031C	SSK2	STE	0.04	78
YNR032W	PPG1	STP	0.05	86
YOL016C	CMK2	CAMK	0.04	84
YOL045W	PSK2	Other	0.03	86
YOL100W	PKH2	AGC	0.04	81
YOL113W	SKM1	STE	0.03	84
YOL128C	YGK3	CMGC	0.04	85
YOR061W	CKA2	Other	0.05	79
YOR090C	PTC5	STP	0.03	84
YOR119C	RIO1		0.03	80
YOR208W	PTP2	PTP	0.04	84
YOR231W	MKK1	STE	0.03	87
YOR233W	KIN4	CAMK	0.04	67
YOR267C	HRK1	Other	0.04	84
YOR351C	MEK1		0.03	84
YPL026C	SKS1		0.05	86
YPL031C	PHO85	CMGC	0.06	80
YPL042C	SSN3	CMGC	0.06	81
YPL140C	MKK2	STE	0.03	72
YPL141C	FRK1	CAMK	0.03	72
YPL150W		CAMK	0.04	85

Table S2 continued.

Kinase or phosphatase	Standard name	Group	FDR	% phosphopeptides*
YPL179W	PPQ1		0.05	82
YPL203W	TPK2	AGC	0.04	83
YPL204W	HRR25	CKI	0.03	78
YPL236C		Other	0.03	80
YPR054W	SMK1	CMGC	0.05	86
YPR073C	LTP1	DSP	0.03	79
YPR106W	ISR1	Yeast PK	0.05	82
YPR111W	DBF20	AGC	0.04	86
YPR161C	SGV1	CMGC	0.03	80

*"Specificity of phosphopeptide enrichment" / "% phosphopeptides": For a definition see Supplementary Materials description of "Database searches" and Bodenmiller *et al*. (*7*). The actual enrichment of phosphopeptides in the sample was likely even higher (for details see the Material and Methods).

Table S4. Significant coregulation of kinases and phosphatases.

Phosphatase	Standard name	Group	Kinase	Standard name	Group	*P* value	Spearman correlation coefficient
YDR075W	PPH3	STP	YIL095W	PRK1	Other	0.00E+00	0.97
YDL230W	PTP1	PTP	YIL042C	PKP1	Atypical PK	0.00E+00	0.95
YNL099C	OCA1		YGR262C	BUD32	Microbial PK	0.00E+00	0.95
YDL230W	PTP1	PTP	YPL026C	SKS1		0.00E+00	0.94
YML057W	CMP2	STP	YAL017W	PSK1	Other	0.00E+00	0.92
YCR079W	PTC6	STP	YGL059W	PKP2	Atypical PK	0.00E+00	0.9
YIR026C	YVH1		YDL025C	RTK1	Other	0.00E+00	0.89
YGR203W	YCH1		YGR262C	BUD32	Microbial PK	0.00E+00	0.86
YNL032W	SIW14	PTP	YDR477W	SNF1	CAMK	1.02E-04	0.85
YML057W	CMP2	STP	YIL042C	PKP1	Atypical PK	2.23E-04	0.87
YLL010C	PSR1		YIL035C	CKA1	Other	2.39E-04	0.88
YLL010C	PSR1		YOR061W	CKA2	Other	2.53E-04	0.71
YDL047W	SIT4	STP	YKL198C	PTK1	Other	4.84E-04	0.84
YML057W	CMP2	STP	YCR008W	SAT4	Other	5.77E-04	0.94
YLR019W	PSR2		YMR139W	RIM11		8.09E-04	0.82
YLL010C	PSR1		YOL045W	PSK2	Other	1.26E-03	0.72
YDR075W	PPH3	STP	YPL203W	TPK2	AGC	1.48E-03	0.75
YLL010C	PSR1		YKL166C	TPK3	AGC	2.04E-03	0.88
YNL099C	OCA1		YFR014C	CMK1	CAMK	4.25E-03	0.88
YDL230W	PTP1	PTP	YAL017W	PSK1	Other	9.45E-03	0.86
YNL099C	OCA1		YMR291W		CAMK	9.87E-03	0.76
YNL099C	OCA1		YIL042C	PKP1	Atypical PK	1.03E-02	0.84
YLR019W	PSR2		YKL171W	NNK1	Yeast PK	1.73E-02	0.98
YDL230W	PTP1	PTP	YDR477W	SNF1	CAMK	2.24E-02	0.76
YML057W	CMP2	STP	YOR231W	MKK1	STE	2.40E-02	0.74
YLR019W	PSR2		YIL042C	PKP1	Atypical PK	2.48E-02	0.88
YLR019W	PSR2		YOL045W	PSK2	Other	2.48E-02	0.88
YDL047W	SIT4	STP	YKL139W	CTK1	CMGC	3.47E-02	0.67
YDL230W	PTP1	PTP	YKL101W	HSL1	CAMK	3.83E-02	0.92
YLL010C	PSR1		YCR008W	SAT4	Other	4.12E-02	0.83
YGR203W	YCH1		YNL307C	MCK1		4.12E-02	0.81
YIR026C	YVH1		YKL166C	TPK3	AGC	4.12E-02	0.81
YNL099C	OCA1		YBR274W	CHK1	Other	4.12E-02	0.78
YLR019W	PSR2		YAL017W	PSK1	Other	4.12E-02	0.78
YDL047W	SIT4	STP	YGR262C	BUD32	Microbial PK	4.12E-02	0.75
YLR019W	PSR2		YFR014C	CMK1	CAMK	4.59E-02	0.85
YLL010C	PSR1		YOR231W	MKK1	STE	4.90E-02	0.73

Table S5. Overlap of data from this study with other data sets.

	All[#]	**Expected direction**		**Inverted direction**	
			Full response		**Full response**
STRING (*28*)	293 (285)**	162**	26**	131	14
SGD (*29*)	170 (168)	93*	22**	77	6
Fiedler (*30*)	39**	28**	4*	11	2
PhosphoGrid (*31*)	16*	10*	3	6	0

*Significant (P <0.05)
**Highly significant (P <0.01)
[#]Given numbers are observed regulatory events, whereas numbers in parentheses indicate the corresponding number of proteins.

Table S7. Overlap of possible direct targets with other data sets.

	All	**Expected direction**		**Inverted direction**	
			Full response		**Full response**
PhosphoGrid *(31)*	3	2	0	1	0
STRING *(28)*	109**	60**	14**	50*	3

*Significant (P <0.05)
**Highly significant (P <0.01)

Table S8. Effects of each kinase and phosphatase on the phosphoproteome. Growth speed is defined according to Hillenmeyer *et al*. (*24*) and phenotypes were defined according to Ohya *et al*. (*23*).

Systematic name	Standard name	Kinase or phosphatase group	Impact rank[1]	Growth phenotype*	Morphology phenotype*
YDL079C	MRK1	CMGC	1		
YML016C	PPZ1	STP	2		
YDR283C	GCN2	Other	4		
YJL187C	SWE1	Other	5		+
YBL056W	PTC3	STP	6		
YGL180W	ATG1	Other	7		
YKL116C	PRR1	CAMK/EMK	8		
YER129W	SAK1		9	+	
YPR073C	LTP1	DSP	10	+	+
YGR040W	KSS1	CMGC	11	+	
YHR076W	PTC7		12		
YDR122W	KIN1	CAMK	13		+
YOR267C	HRK1	Other	14		
YKL048C	ELM1		15		++
YLR248W	RCK2		16	+	
YOL016C	CMK2	CAMK	17		
YCR091W	KIN82	AGC	18	+	
YBR028C		AGC	19		+
YPL236C		Other	20		
YNL161W	CBK1	AGC	21		
YIL113W	SDP1		22		
YPL150W		CAMK	23		
YHR082C	KSP1		24		+
YOR351C	MEK1		25		
YDR247W	VHS1	Other	26	+	
YOL113W	SKM1	STE	27		+
YBL009W	ALK2		28		
YLR362W	STE11	STE	29		
YDL028C	MPS1	Other	30		
YDR523C	SPS1	STE	31		++
YDL101C	DUN1		32		
YPL141C	FRK1	CAMK	33		
YPR161C	SGV1	CMGC	34		

[1]Highest impact = 124, lowest impact = 1. *Compared to wild-type; "+", strong; "++", very strong (see Material and Methods).

Table S8 continued.

Systematic name	Standard name	Kinase or phosphatase group	Impact rank	Growth phenotype*	Morphology phenotype*
YGL179C	TOS3	CAMK	35		
YOR090C	PTC5	STP	36		+
YOL128C	YGK3	CMGC	37		
YKL161C	KDX1	CMGC	38		+
YKL168C	KKQ8	Other	39		+
YPL204W	HRR25	CKI	40		
YOR119C	RIO1		41		
YBL088C	TEL1	Inositol Kinase	42		+
YLR113W	HOG1	CMGC	43	+	+
YJL128C	PBS2	STE	44		+
YGR123C	PPT1	STP	45		
YBL016W	FUS3	CMGC	46		++
YJL095W	BCK1	STE	47	++	+
YPR106W	ISR1	Yeast PK	48	+	
YPL179W	PPQ1		49		
YHR030C	SLT2	CMGC	50	++	+
YDL214C	PRR2	Other	51		
YJL106W	IME2	CMGC	52		
YDL006W	PTC1	STP	53		++
YMR104C	YPK2	AGC	54		+
YGL021W	ALK1		55		
YPR054W	SMK1	CMGC	56	+	
YAR019C	CDC15	STE	57		
YOR233W	KIN4	CAMK	58		
YOR208W	PTP2	PTP	59	+	
YBR059C	AKL1	Other	60	+	
YNL020C	ARK1	Other	61		+
YBR160W	CDC28	CMGC	62		
YNR032W	PPG1	STP	63		+
YJR059W	PTK2	Other	64	+	
YLR096W	KIN2	CAMK	65		
YJL164C	TPK1	AGC	66		+
YGR188C	BUB1	Other	67		
YBR097W	VPS15	Other	68		++

[1] Highest impact = 124, lowest impact = 1. *Compared to wild type; "+", strong; "++", very strong (see Material and Methods).

Table S8 continued.

Systematic name	Standard name	Kinase or phosphatase group	Impact rank	Growth phenotype*	Morphology phenotype*
YPR111W	DBF20	AGC	69		
YDL108W	KIN28	CMGC	70		
YMR139W	RIM11		71		
YNL032W	SIW14	PTP	72	+	+
YBR125C	PTC4	STP	73		
YJL165C	HAL5	Other	74	++	
YDL025C	RTK1	Other	75		
YNR031C	SSK2	STE	76		+
YLL019C	KNS1	CMGC	77		
YLR019W	PSR2		78		
YBR274W	CHK1	Other	79		
YMR216C	SKY1	CMGC	80	++	++
YPL140C	MKK2	STE	81		
YFR014C	CMK1	CAMK	82		
YJR066W	TOR1	Inositol Kinase	83		
YIL095W	PRK1	Other	84		
YOL100W	PKH2	AGC	85		
YNL298W	CLA4	STE	86		++
YDR466W	PKH3	AGC	87		+
YMR291W		CAMK	88		
YGL059W	PKP2	Atypical PK	89		
YDR075W	PPH3	STP	90	++	
YLR433C	CNA1	STP	91		
YBR276C	PPS1		92		+
YML057W	CMP2	STP	93		
YDL188C	PPH22	STP	94		
YCR008W	SAT4	Other	95	+	
YNL099C	OCA1		96	++	
YCR079W	PTC6	STP	97	+	
YLR240W	VPS34	Inositol Kinase	98		+
YIL035C	CKA1	Other	99	+	+
YPL031C	PHO85	CMGC	100		+
YPL026C	SKS1		101		+
YOR061W	CKA2	Other	102	+	++

[1] Highest impact = 124, lowest impact = 1. * compared to wild type; "+" strong, "++" very strong (See Material and Methods).

Table S8 continued.

Systematic name	Standard name	Kinase or phosphatase group	Impact rank	Growth phenotype*	Morphology phenotype*
YOR231W	MKK1	STE	103		++
YDL134C	PPH21	STP	104		
YGR203W	YCH1		105		
YKL101W	HSL1	CAMK	106		++
YAL017W	PSK1	Other	107		
YKL171W	NNK1	Yeast PK	108		+
YKL166C	TPK3	AGC	109		
YHR135C	YCK1	CKI	110	+	
YIL042C	PKP1	Atypical PK	111		+
YDL230W	PTP1	PTP	112		
YPL203W	TPK2	AGC	113		
YNL307C	MCK1		114	++	+
YIR026C	YVH1		115	+	++
YGR262C	BUD32	Microbial PK	116		++
YKL126W	YPK1	AGC	117	++	++
YPL042C	SSN3	CMGC	118		++
YDL047W	SIT4	STP	119		++
YLL010C	PSR1		120		+
YDR477W	SNF1	CAMK	121	++	
YKL198C	PTK1	Other	122		
YOL045W	PSK2	Other	123	+	
YKL139W	CTK1	CMGC	124		++

[1]Highest impact = 124; lowest impact = 1.

*Compared to wild-type; "+", strong; "++", very strong (see Material and Methods).

Table S9. Enrichment of biological process among the low-impact kinases (bottom half). n.d., not determined.

Biological process	***P* value**	**Corresponding *P* value for bottom half kinases**
protein kinase cascade	1.9E-11	5.8E-05
MAPKKK cascade	3.9E-10	8.6E-04
osmosensory signaling pathway	5.6E-10	1.8E-01
cell surface receptor linked signal transduction	1.8E-09	9.8E-02
primary metabolic process	3.4E-09	3.0E-07
cellular metabolic process	1.1E-08	8.3E-06
MAPKKK cascade during osmolarity sensing	2.4E-08	9.6E-02
metabolic process	5.9E-08	3.4E-05
regulation of MAP kinase activity	9.3E-08	n.d.
response to osmotic stress	1.2E-07	5.8E-01
regulation of cell cycle	2.4E-07	1.2E-04
regulation of molecular function	2.6E-07	1.3E-02
regulation of protein kinase activity	4.9E-07	2.7E-01
intracellular signaling cascade	5.6E-07	5.6E-07
regulation of kinase activity	5.9E-07	2.8E-01
regulation of conjugation	7.1E-07	2.8E-01
regulation of conjugation with cellular fusion	7.1E-07	2.8E-01
regulation of multi-organism process	7.1E-07	2.8E-01
regulation of transferase activity	8.4E-07	2.9E-01
regulation of catalytic activity	1.5E-06	5.1E-02
response to stimulus	1.9E-06	4.7E-07
regulation of cell division	2.1E-06	3.3E-01
response to pheromone	4.2E-06	7.1E-02
cell cycle	5.5E-06	5.5E-06
regulation of cellular component organization and biogenesis	6.4E-06	5.0E-03
inactivation of MAPK activity during osmolarity sensing	7.2E-06	n.d.
inactivation of MAPK activity	7.2E-06	n.d.
protein amino acid autophosphorylation	7.2E-06	4.5E-02
negative regulation of MAP kinase activity	7.2E-06	n.d.
negative regulation of protein kinase activity	1.4E-05	n.d.
negative regulation of transcription by pheromones	1.4E-05	n.d.
negative regulation of transcription from RNA polymerase II promoter by pheromones	1.4E-05	n.d.

Biological process	***P* value**	**Corresponding *P* value for bottom half kinases**
cell cycle checkpoint	1.7E-05	2.2E-04
response to stress	1.8E-05	1.8E-05
negative regulation of kinase activity	2.5E-05	n.d.
regulation of cell morphogenesis	3.3E-05	1.6E-03
regulation of anatomical structure morphogenesis	3.3E-05	1.6E-03
negative regulation of transferase activity	3.9E-05	n.d.
regulation of developmental process	4.9E-05	2.1E-03
cell division	5.5E-05	8.8E-03
negative regulation of specific transcription from RNA polymerase II promoter	5.9E-05	n.d.
negative regulation of gene-specific transcription	5.9E-05	n.d.
regulation of meiosis	6.9E-05	n.d.
regulation of meiotic cell cycle	8.1E-05	n.d.
negative regulation of conjugation	8.2E-05	n.d.
negative regulation of conjugation with cellular fusion	8.2E-05	n.d.
negative regulation of multi-organism process	8.2E-05	n.d.
regulation of transcription by pheromones	8.3E-05	n.d.
regulation of transcription from RNA polymerase II promoter by pheromones	8.3E-05	n.d.

Table S10. Enrichment of biological process among the high-impact kinases (top half). n.d., not determined.

Biological process	***P* value**	**Corresponding *P* value for top half kinases**
interphase of mitotic cell cycle	3.3E-09	2.7E-01
G1/S transition of mitotic cell cycle	4.9E-09	3.4E-01
interphase	5.2E-09	2.9E-01
primary metabolic process	3.0E-07	3.4E-09
response to stimulus	4.7E-07	1.9E-06
intracellular signaling cascade	5.6E-07	5.6E-07
peptidyl-serine phosphorylation	7.3E-07	n.d.
peptidyl-serine modification	7.3E-07	n.d.
cell cycle phase	1.2E-06	1.8E-02
mitotic cell cycle	1.4E-06	8.7E-02
cell cycle process	4.1E-06	3.0E-02
cell cycle	5.5E-06	5.5E-06
cellular metabolic process	8.3E-06	1.1E-08
cellular ion homeostasis	1.1E-05	3.0E-01
cellular chemical homeostasis	1.1E-05	3.0E-01
ion homeostasis	1.6E-05	3.2E-01
chemical homeostasis	1.7E-05	3.2E-01
response to stress	1.8E-05	1.8E-05
regulation of biological quality	2.1E-05	4.8E-03
cellular developmental process	2.1E-05	3.9E-04
cell morphogenesis	2.2E-05	1.2E-04
anatomical structure morphogenesis	2.2E-05	1.2E-04
cellular structure morphogenesis	2.2E-05	1.2E-04
anatomical structure development	2.2E-05	1.2E-04
metabolic process	3.4E-05	5.9E-08
protein kinase cascade	5.8E-05	1.9E-11
regulation of mitotic cell cycle	6.0E-05	1.5E-01

Table S12. Overview of the entire data set.

	Total number	Regulated	Not regulated
Phosphoproteins	1,677 [1]	1,029 [2]	648 [3]
Phosphopeptides	11,374 [4]	3,824 [5]	7,550 [6]
Phospho-Events	158,168 [7]	8,814 [8]	149,354 [9]

[1]Number of all identified phosphoproteins in the dataset. [2]Number of all phosphoproteins that had at least one regulated phosphorylation site. [3]Number of all phosphoproteins that did not have any regulated phosphorylation site. [4]Number of all identified phosphopeptides in the dataset. [5]Number of all phosphopeptides that were considered to be regulated. [6]Number of all phosphopeptides that were not regulated. [7]Number of all individually identified phosphorylation sites per kinase or phosphatase. [8]Number of all individually identified regulated phosphorylation sites per kinase or phosphatase. [9]Number of all individually identified phosphorylation sites that were not regulated per kinase or phosphatase.

8

ADDITIONAL RESULTS AND DISCUSSION

As discussed in [83], our analysis revealed, among other things, two intriguing observations: a) a surprisingly large number of detected phosphopeptides that never exhibited a significant change in abundance upon the knockout of any given kinase or phosphatase, and b) a high prevalence of indirect effects, as evidenced by the large number of peptides peptides whose abundance changes had a directionality opposite to the expectation. In our work we discussed possible reasons for these findings: i) conservative cutoffs might have prevented the detection of a subset of events, ii) a number of essential kinase and phosphatase mutants could not be analyzed, iii) paralogous kinases and phosphatases possibly have redundant functions and consequently may have masked each other, and last but not least iv) the experimental set-up itself may have been too restricted.

However, it is worth mentioning that the characteristics discovered in the yeast phosphorylation network meet the requirements for a comprehensive regulatory cellular signaling network: interconnectedness and dynamic flexibility are high, while all the same a stable cellular state is maintained against most of our perturbations [83, 84, 85, 86].

These observed network traits are associated with the concept of biological robustness, one of the fundamental properties of all biological systems [87, 88]. The network's interconnectivity and non–linearity in terms of its large number of connections between kinases, phosphatases and phosphoproteins (which are often kinases themselves) form the basis for this robustness: first, most kinases and/or phosphatases have an extensive impact on the phosphoproteome, affecting a wide range of biological processes [83] and second, the network comprises a large number of indirect effects. Furthermore, as mentioned above, phosphopeptides failing to appear as regulated could indicate a functional redundancy in the network where similar kinases and phosphatases can replace each other. Due to the network size and its interconnectivity, one could also imagine that the (de)phosphorylation of a peptide is backed up by alternative network pathways not functionally related to the particular perturbed kinase or phosphatase. Either mechanism contributes to the network's robustness by buffering perturbations [87, 88]. This also substantiates the observation that in most cases, no strong cellular phenotypes were detectable in our single knockout mutants [83].

As described above, a certain degree of indirectness in the phosphorylation network was to be expected. Nevertheless, the lack of sufficient direct

kinase or phosphatase targets naturally complicated the analysis of our data.

For instance, we grouped potential direct molecular phosphopeptide targets according to their responsible kinase and scanned each set of phosphopeptides for overrepresented patterns using the motif–x software [89]. Then, we compared these patterns to known consensus recognition sites of the respective kinase [90]. However, none of the kinase phosphorylation site motifs could be recovered. In a second step, we restricted the set of possible direct kinase targets to phosphopeptides that showed a 'full response' (vanisher peptides). This resulted in too little phosphopeptides per kinase, preventing statistically significant conclusions. For this reason, we finally clustered the peptides according to whole kinase families (see Figure 8(a)) instead to single kinases. Figure 6 shows the results for consensus recognition sites of three kinase families.

Except for the sequence logo derived from possible direct phosphopeptide targets of the CAMKL kinase family, none of the logos revealed any evidence for a specific kinase recognition signal. In addition, there were no apparent differences in the expressiveness of sequence logos derived from direct and indirect target sets. Furthermore, the direct target sequence logo of the CAMKL family could not be identified in the set of known consensus phosphorylation site motifs of kinases in this family. These negative results are most likely due to the limited number of direct targets in our dataset, and the analysis strongly supports our finding of a high number of indirect effects in our dataset.

The clustering of kinases while using different measures of similarities was another important aspect in our study. For instance, we analyzed all kinases with regard to their effect on the phosphoproteome and searched for common properties between kinases showing similar phosphorylation 'footprints'. In doing so, the similarity between pairs of kinases was determined using different measurements.

One approach was to simply count the number of phosphopeptides which are affected by both kinases. Figure 7 relates this 'shared–substrate' similarity between kinase pairs to the type of genetic interaction they exhibit. As can be seen, there is a weak but significant correlation: the more phosphopeptides two kinases share, the more likely they also interact genetically, suggesting that kinases that act on the same peptides are also functionally related. Notably, this result is consistent with the findings presented in [69].

Comparing kinase similarity to other properties like the number of PPI shared by the kinases, or the number of gene ontology (GO)–terms shared, did not yield significant results (data not shown).

A more sophisticated similarity measure for kinases is the correlation coefficient between the fold changes of the phosphopeptides that are affected by the given kinases. For example, we performed a clustering ana-

lysis based on the Spearman correlation of the kinases' phosphosite fold change vectors and reviewed the clusters obtained for a) reflecting the kinase family tree [29] (Figure 8(a)) and b) being consistent with functional annotations of the kinases, for example regarding their GO annotations [92]. In both cases no significant enrichment could be found.

Similarity between kinases can also be defined based on characteristics other than their measurable effects on the phosphoproteome. For instance, pairs of kinases working together in pathways or functional units should also exhibit similarities in terms of their cellular and morphological phenotypes. Just as before, we used correlation analysis to determine the similarity of kinases with respect to their cellular phenotypic profiles (see Table S8 in Supplementary Methods).

By invoking both the similarities concerning the phosphoproteomic phenotype and the cellular phenotype we carried out a bi–clustering analysis in order to test whether kinases that have similar phosphoproteome footprints also tend to have similar phenotypic consequences on the cell. Furthermore, we included a third source of information by mapping all kinases to their respective kinase superfamily. The result is shown in Figure 8(b).

As can be seen, there is no obvious relationship between similar phosphoproteomic phenotypes and similar cellular phenotypes. Although some of the kinase pairs cluster together in both phenotypic dimensions (red edges), the majority of the pairs displays an independent behavior and did not support our hypothesis that kinases that share similar effects on substrates also tend to have similar morphological effects.

Besides the clustering of kinases, we also clustered all regulated phosphopeptides according to their kinase or phosphatase regulation profiles. On the one hand we screened for functional uniformity of the clusters, on the other hand for peptides of a cluster being enriched in similar kinase binding motifs. Again, the results were largely uninformative.

Bearing in mind that a number of kinases and phosphatases could not be analyzed for technical reasons and the existence of intrinsic technical shortcomings within the experimental setup, we expected that our data might not be representative of the whole regulated phosphoproteome. Nevertheless, the restricted overlap of our data with known kinase–substrate interactions as listed in various PPI databases [94, 95] was dissatisfying. However, to some extent this could be explained by the transient type of interaction between kinases and phosphatases and their substrates, which common PPI databases likely fail to specify. Furthermore, PhosphoGrid [15], the only database focusing on kinase/phosphatase – substrate interactions, currently includes a rather small portion of the phosphoregulome.

Another important aspect must be taken into account with regard to a) the indirect and incomplete nature of the data and b) the amount of non–regulated phosphopeptides: In order to be considered as a possible direct

target of a kinase, a phosphopeptide would have to show either a decrease in abundance or a complete absence in the knockout sample of that kinase. This necessarily requires the successful detection of the phosphopeptide in the corresponding wild–type samples (to which the mutant samples are compared to). Otherwise, the peptide only appears to be phosphorylated in the mutant and may consequently be considered as an inverse responder.

In our dataset, a total of 102 wild–type samples (34 batches of wild–type cells with 3 replicates each) were measured. In contrast, 372 mutant samples, more than 3.5 times more, were processed. This indicates on the one hand a notable undersampling of the wild type relative to the mutants and moreover, it means that some of the wild–type batches were compared to several deletion strains. In theory, one would expect different perturbed phosphoproteome maps aligned to the same wild–type samples to offer the same wild–type phosphopeptides. However, close scrutiny revealed that the number of identified peptides varies within a range of 30%, an inconsistency which most likely prevented the identification of direct targets in many cases and may sometimes even have favored the inverse conclusion.

Another test to reveal deficiencies concerning the wild–type reference sets is to look for kinases and phosphatases that are represented in the data by their own phosphopeptides. A kinase or phosphatase that had been knocked out (i.e. genetically deleted) should not be detectable at all in the corresponding mutant sample. Therefore, possible self–phosphopeptides should disappear and be reported as 'vanisher' peptides. Such peptides should act as a control set. We checked our dataset and found 58 kinases that were indeed detected at least once as phosphopeptides. However, only for 7 of them (12 %) the expected vanisher events were observed, indicating an under–reporting of about 90% of this most obvious class of vanisher events. On the positive side, only a single peptide was wrongly called as present (i.e., a false positive) in a sample where the corresponding kinase gene had been deleted.

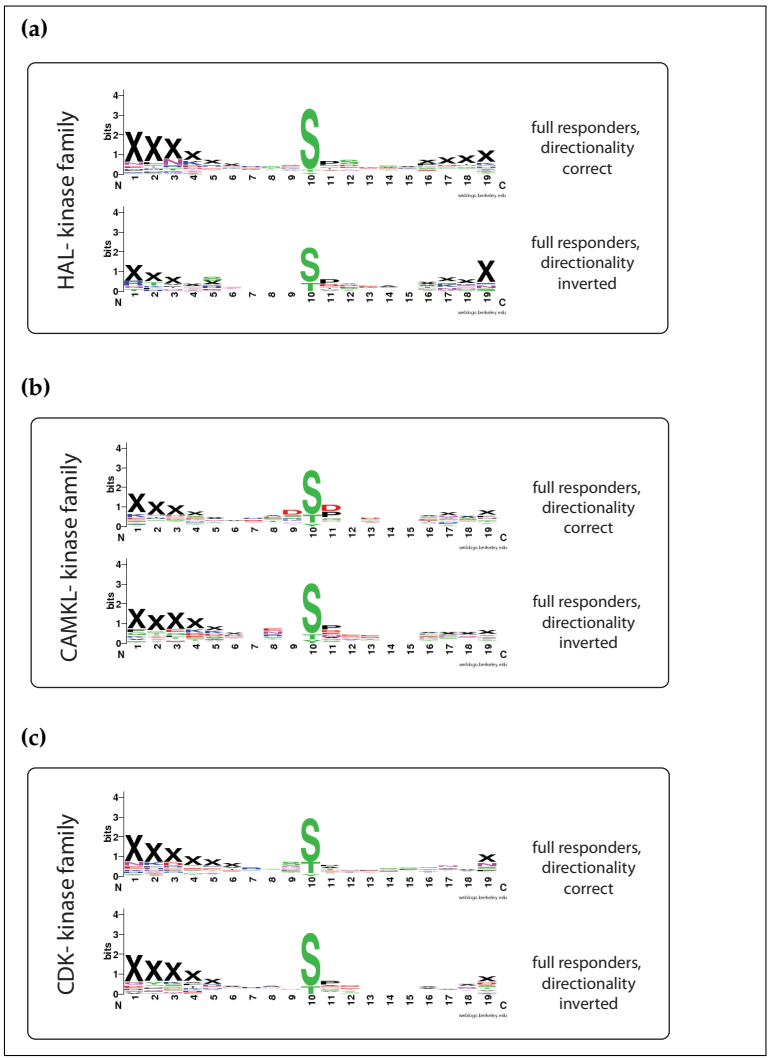


Figure 6: *Consensus phosphorylation site motifs.* Full responder phosphopeptides with expected directionality were grouped according to their kinases' families and scanned for overrepresented consensus recognition sites. For a negative control, the analysis was repeated for strong indirect targets, i.e. full responders with inverted directionality. **(a)** Obtained sequence logo for kinase family HAL, both for the set of direct and indirect targets. **(b)** Sequence logo for kinase family CAMKL. **(c)** Sequence logo for kinase family CDK.

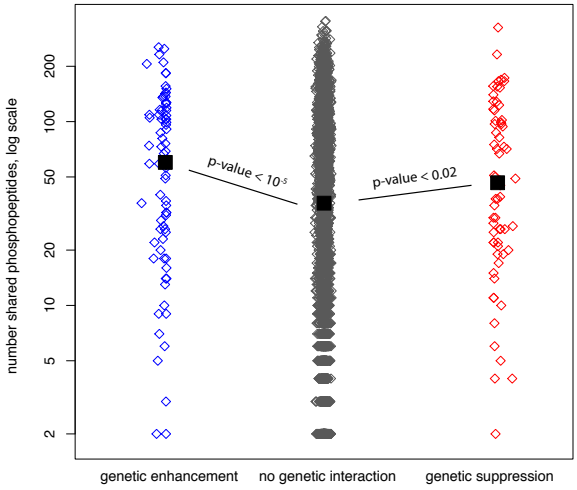


Figure 7: *Similarity of kinase pairs versus genetic interactions*. Each diamond represents a pair of kinases. Pairs that show a positive genetic interaction are plotted in blue, pairs with negative genetic interaction are plotted in red, and pairs that are not known to interact genetically are plotted in grey. Black rectangles represent the median number of shared phosphopeptides. Y–axis: number of phosphopeptides shared between two kinases, in log space. The genetic interaction data was taken from [91].

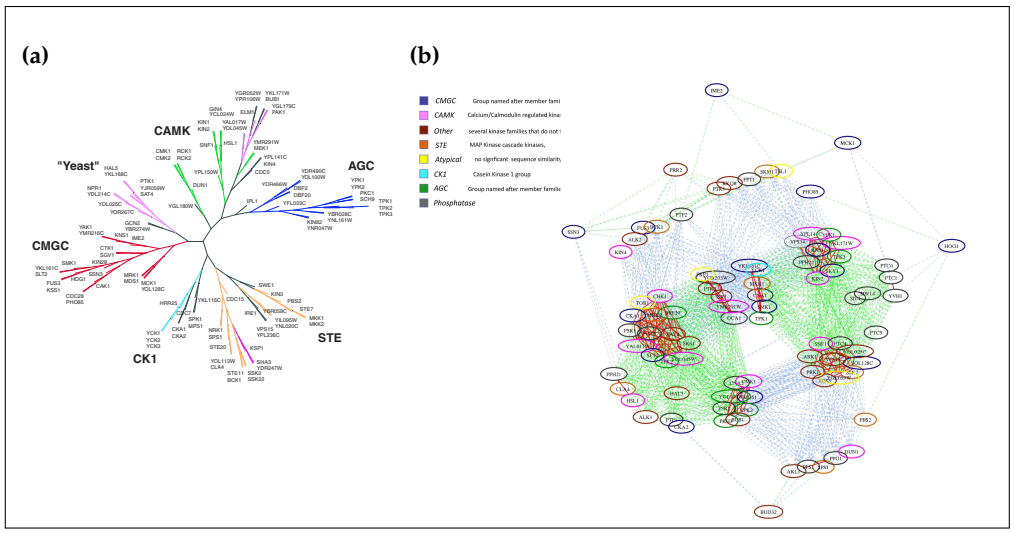


Figure 8: *Bi–clustering analysis of kinases* **(a)** Dendrogram of the yeast protein kinase superfamilies. Taken from [93]. **(b)** Clustering network obtained from the simultaneous clustering of kinases based on both their phosphosite response similarity and their morphological similarity. Kinases are represented as nodes in the network and are colored according to the their superfamily membership. Edge colors: i) red: corresponding kinases are grouped together with respect to both dimensions, ii) green: corresponding kinases only cluster with respect to the similarity of their phosphoproteomic phenotypes, iii) blue: corresponding kinases only cluster with respect to the similarity of their morphological phenotypes.

9

FURTHER PERSPECTIVES

To overcome the inconsistencies mentioned above, we intend to merge all wild–type samples simultaneously in order to create a global 'super'–reference map. This will allow us to aggregate all phosphopeptide information contained in the wild types ('artificial deep sampling') and ensures that all mutants are compared to the same list of wild–type phosphopeptides.

Furthermore, we plan to use different LC–MS/MS alignment software programs like MaxQuant [64] and OpenMS [60] in order to compensate for the various strengths and weaknesses of the individual programs. The ultimate goal will be to generate a 'gold standard' of phosphopeptides that were identified as regulated while using different comparative alignment software programs.

Part III

PHOSPHOPROTEOMICS AND METABOLOMICS IN SYSTEMS BIOLOGY

10 INTRODUCTION

10.1 METABOLOMICS

Metabolites are small molecules that take part in or are produced during various metabolic processes. They do not only comprise a very diverse spectrum of chemical compounds including lipids, carbohydrates, and amino acids, but are also present in a wide range of concentrations. The complete set of all small molecules in a defined system (e.g. a cell or tissue) forms the metabolome [96, 97]. Estimates of metabolome sizes differ greatly; for example, predictions for the plant kingdom range from 200.000 to 1.000.000 metabolites [98]. As for the human metabolome, currently about 8000 curated metabolites are known but estimates go up to 30.000 and more [99]. For yeast, a recent study raised the number of described metabolites to about 950 [100].

In a manner analogous to transcriptomics and proteomics, *metabolomics* describes the global study of the metabolome from a Systems Biology approach. Other terms often used in this context are metabonomics, metabolic profiling, and metabolic fingerprinting. Metabolomics continues to be a rapidly developing field and terminologies evolve constantly, resulting in ongoing debates within the scientific community about the respective definitions (which also often overlap) [97, 101]. However, in recent years *metabolomics* has emerged as the term used most often [102, 103].

Like with other 'omics'–fields technological developments have been (and are) the basic requirements for system–wide studies of the metabolome. Primarily the use of two HTP procedures, namely nuclear magnetic resonance (NMR) and MS, has enabled the quantification of large numbers of metabolites in a non–targeted manner [104]. NMR has the advantage of being a noninvasive platform with relatively low requirements in terms of sample preparation. While having an excellent quantitative precision the limiting step of NMR is its sensitivity, allowing only the detection of metabolites with medium to high abundances. In contrast, the detection ability of MS techniques covers a broad range of compound concentrations. While separation–based techniques like LC–MS and gas chromatography mass spectrometry (GC–MS) predominated in the beginning, recently (direct) flow injection electrospray mass spectrometry (FIE–MS) has been established, offering a more 'global' coverage of the metabolome [104, 103, 105, 106].

Despite substantial technical advances, two main challenges persist: first, it is still not feasible to identify all metabolites in a sample, and second, the

quantification of detected compounds is limited even further. Most studies barely achieve an identification rate of 25% [97, 107].

System–wide metabolic studies assist in understanding metabolism in general, for instance by predicting novel metabolic pathways and by providing input for metabolic network construction [108, 109, 110]. In functional genomics, comparative metabolomics is employed for elucidating enzyme functions or in identifying natural substrates of enzymes [104, 111].

Scientifically, applications for metabolomics are versatile. One prominent and significant field is medical sciences. For example, the discovery of new biomarkers plays an important role in clinical diagnosis. The number of diagnostic assays using metabolic markers is increasing, amongst others signature studies for breast cancer [112], diabetes [113], and Alzheimer's disease [114] have been reported over the past years. Furthermore, another important field of application is pharmacology. Metabolomic studies are particularly suitable for toxicology and pharmacogenomics [115, 103]. Ultimately, the goal is to employ metabolomics as a tool to pursue and establish personalized medicine [116, 117].

10.2 REGULATION OF THE METABOLOME

The metabolome represents a fairly downstream component of the 'omics' cascade [118]. Consequently, mechanisms controlling the upper–ome levels such as the transcriptional and the phosphorylation regulatory networks have a strong influence on the metabolic network, transmitted through the enzymes that are catalyzing the metabolic reactions. This allows for tight regulation, and as a matter of fact, transcriptional and PTMs are important factors in metabolic control (see Figure 9). Also referred to as *hierarchical*, these types of control affect the quantity of active enzymes whereas *metabolic* regulation targets the enzymes' kinetic activities by means of the presence of substrates, products, and modifiers [119, 120].

Transcriptional regulation of the metabolism has been a main focus of study for many years. With the emergence of transcriptomics, genome–wide screens could be performed in order to analyze the comprehensive effect of changes in abundance of particular enzyme transcripts on metabolic networks in a variety of organisms [121, 119, 122, 123].

In contrast, the global impact of PTMs on metabolic control is less well understood. Concerning protein acetylation, two recent reports suggest a more relevant role than previously thought [124, 125]. However, the key player in post–translational regulation appears to be the process of reversible protein phosphorylation [126, 127]. While we know of some specific examples of metabolic enzymes regulated by protein phosphorylation, to the best of my knowledge, no system–wide studies have been published thus far. Even though great progress has been made towards cataloguing and

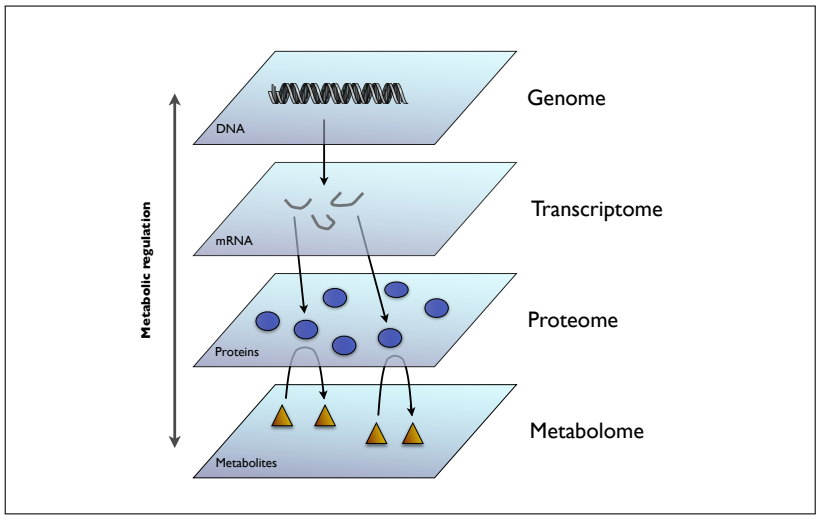


Figure 9: *The 'omics' cascade.* The hierarchically coordinated 'omics' cascade comprises complex datasets that interact in highly concerted actions. The metabolome is one of the endpoints of the cascade and thus perhaps the closest to the phenotype. Adapted from [118].

mapping phosphorylation sites, the "function of specific phosphorylation events has rarely been demonstrated and quantified, leaving a major challenge for the near feature" [126], and it "remains to be established whether those modifications actually have regulatory significance for protein function or not" [128].

Nonetheless, metabolic regulation by protein phosphorylation and regulation by transcription are closely connected. For example, many TFs are phosphorylated in order to perform transcriptional initiation. Consequently, as illustrated in Figure 10, the phosphorylation network and the transcriptional network are highly intertwined, and a direct (and isolated) analysis of the functional role of protein phosphorylation on the metabolome is thus hardly possible.

Below I provide some insight into the first system–wide study on the impact of reversible protein phosphorylation on metabolic regulation in yeast. This project, conducted in close collaboration with Juliane Schulz from the Uwe Sauer group, ETH Zurich, is a work in progress.

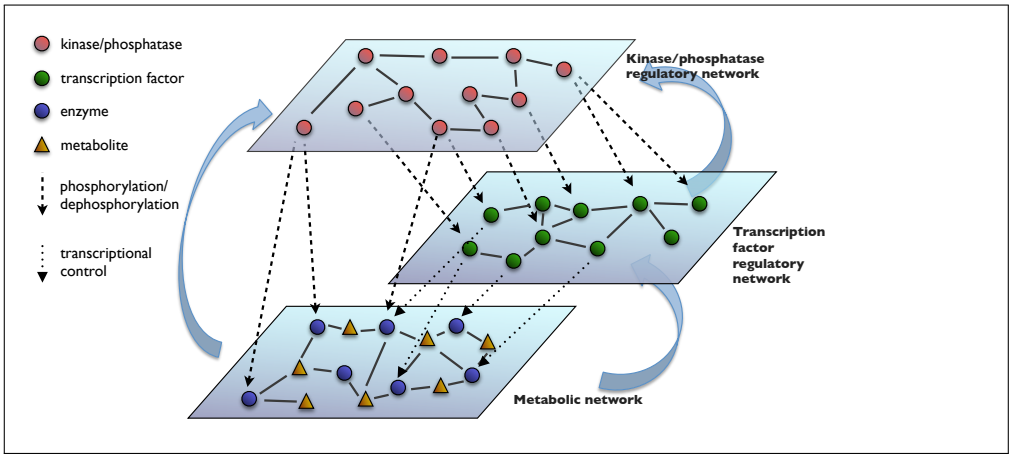


Figure 10: *Interconnectedness of the phosphorylation network, transcription factor network, and metabolic network.* Reversible protein phosphorylation controls both the TF network and the metabolic network, the latter being coordinated via transcriptional control, too. In contrast, via feedback loops, metabolic levels also may influence the regulatory networks.

11

INTEGRATING THE PHOSPHOPROTEOME AND THE METABOLOME

11.1 PROJECT DESCRIPTION

As highlighted in the previous section, little is known about the control of metabolism by protein phosphorylation on a global scale. One approach to address this is to analyze the set of proteins that execute this very type of PTM: protein kinases and protein phosphatases. Their systematic silencing coupled to subsequent quantitative metabolomics analysis yields metabolites whose abundances are affected (either directly or indirectly) by the activity of a given kinase or phosphatase. Consequently, these metabolites are potential candidates for regulation via protein phosphorylation.

However, exactly *how* these metabolites are linked to the kinase and phosphatase regulatory network remains to be understood. Does their abundance depend on the phosphorylation of distinct metabolic enzymes? More precisely, can the behavior of particular metabolites be associated with individual phosphosites? Or could these variations in metabolite concentration perhaps be a result of changes in the transcriptional level of enzymes, caused by phosphorylation–based activation or inactivation of TFs?

The availability of datasets from kinase and/or phosphatase perturbation experiments both on the phosphoproteome [83] and on the transcriptome [71] level provides an opportunity to address these questions. The integrated analysis of the datasets can help to shed light on the complex relationships between the three '–omics' levels and to elucidate regulation of the metabolome by protein phosphorylation.

Implementing the experimental approach mentioned in the first paragraph, the group of Uwe Sauer at ETHZ generated 122 perturbed metabolomes by performing FIE–MS [106] on the same knockout mutant strains that were used in the genome–wide kinase and phosphatase knockout phosphoproteome screens (91 kinases, 27 phosphatases) in the yeast *Saccharomyces cerevisiae* described before [83]. Significantly and reproducibly regulated ions were detected by comparing the abundance of mutant metabolomes to those of wild–type metabolomes [129].

The data obtained – in combination with the phosphoproteomic data described in chapter 7 – enables to address the following questions:

- To what extent does the phosphorylation network regulate the metabolomic network?

- Which are the specific kinases and phosphatases involved in metabolic regulation?
- Which metabolic enzymes and metabolic pathways are regulated by protein phosphorylation?
- Can we find metabolite–phosphosite pairs that show a correlated behavior in their abundance changes upon each knockout of the same set of kinases and phosphatases?

11.2 RESULTS AND DISCUSSION

In total, 1710 distinct ions could be quantified. Ions exceeding a $\log_2$ fold change of 0.8 (q-value $\leqslant 0.001$, FDR–based multiple testing corrected) were defined as regulated, resulting in a total of 662 ions (39%). The number of kinases and phosphatases affecting a given ion ranged from 1 to 81, with an average number of five. Only four ions had a regulatory link to more than 70 kinases or phosphatases. About 415 of all ions could be annotated and assigned to chemical compounds (24.3%).

Of the 91 kinases and 27 phosphatases tested, 78 and 26, respectively, showed an effect on the metabolome, that is they appeared to regulate at least one ion. On average, a kinase or phosphatase regulated 34 ions; six kinases (Vps34, Pho85, Vps15, Ime2, Snf1, Ctk1) and one phosphatase (Ptc4) affected more than 100 ions.

In order to find pairs of metabolites and phosphorylation sites that responded similarly to the absence of a set of kinases and phosphatases, we calculated the correlation (both Pearson and Spearman rank correlation) between the phosphosite and the metabolite abundance fold changes. In doing so, phosphosite and metabolite both had to be showing an effect in at least five shared kinase or phosphatase perturbations. To increase the number of possible pairs we also considered fold changes (for both metabolites and phosphosites) that fell below the original regulation cutoffs (phosphosite: $|\log_2 \text{FC}| > 1.5$, p-value $\leqslant 0.015$, metabolite: $|\log_2 \text{FC}| > 0.8$, q-value $\leqslant 0.001$).

Figure 11 compares the number of actually observed significantly correlated metabolite–phosphosite pairs (red line, correlation p-value $\leqslant 0.01$) to the expected number of significant pairs (box plots, based on 1000 randomizations), plotted as a function of the significance cutoff of the corresponding metabolite and phosphosite fold changes (*data p-value*, x–axis). The individual plots vary in the required minimum number of kinases and/or phosphatases that each phosphosite–metabolite pair had to share in order to be considered for the correlation analysis. (Figure 11(a): 5, Figure 11(b): 10, Figure 11(c): 15, Figure 11(d): 20).

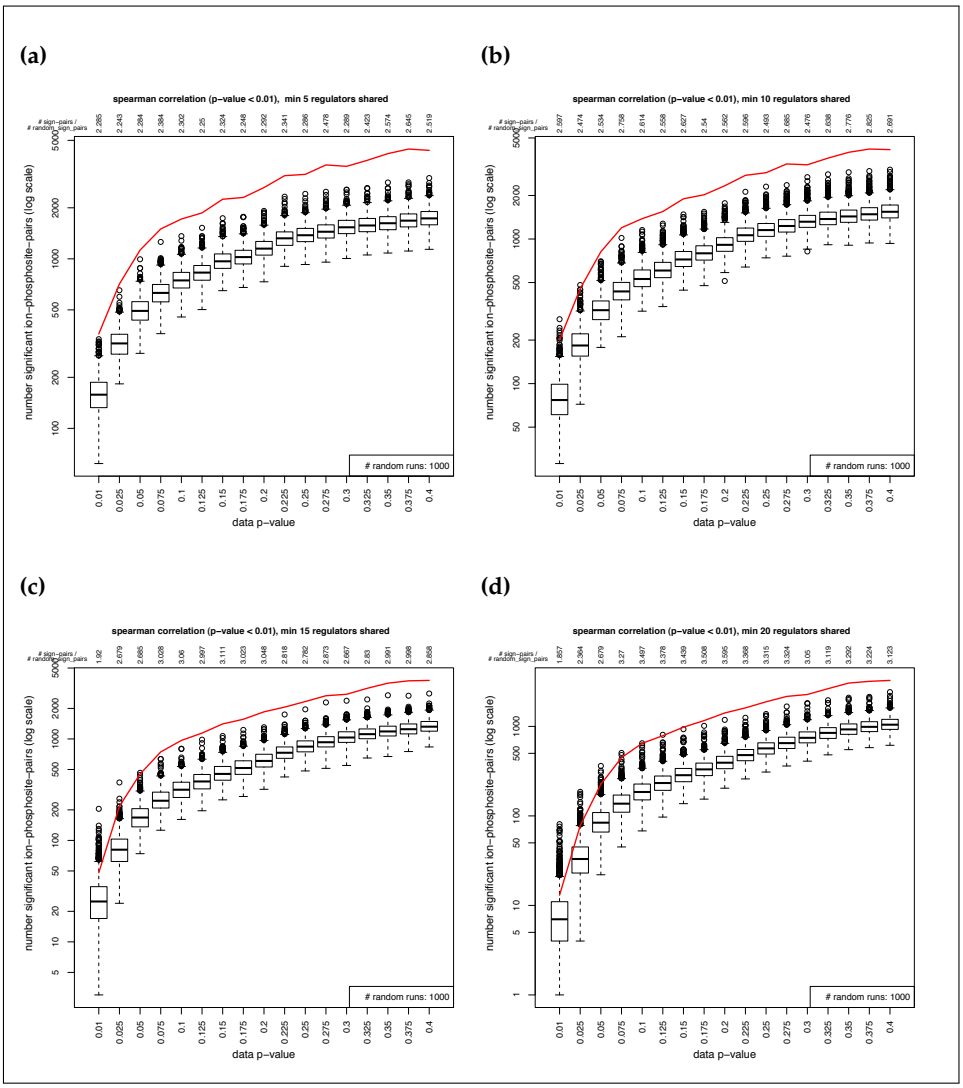


Figure 11: *Number of observed significantly correlated metabolite–phosphosite pairs versus the expected number of significant pairs.* The number of expected pairs was estimated by calculating the correlations for metabolite–phosphosite pairs with randomized fold changes and repeating this procedure for 1000 times. The fold change randomization was done by shuffling the links between phosphosites and kinases/phosphatases in the phosphoproteome dataset. The metabolomics data was kept unchanged. The plots differ in the requested minimum number of shared kinases/phosphatases. **(a)**: minimum of 5 shared kinases/phosphatases. **(b)**: minimum of 10 shared kinases/phosphatases. **(c)**: minimum of 15 shared kinases/phosphatases. **(d)**: minimum of 20 shared kinases/phosphatases.

In all plots, the number of significantly correlated phosphosite–metabolite pairs (y–axis, log space) is clearly above the random expectation, which is encouraging for our analysis. Relaxing the individual fold change cutoffs (x–axis) is likewise justified: pairs whose correlation was calculated based on fold changes that were not necessarily considered to be significantly regulated (x–axis, data p-value $\geqslant 0.015$) individually, also clearly outnumber the expected number of pairs. This outcome suggests to expand the cutoffs for further analyses, too.

Across the whole range of data p-value cutoffs, except for the first one of 0.01, the ratio of the number of observed significant data pairs to the number of expected significant data pairs shows only very little variation. In contrast, more stringent conditions in terms of the minimum number of shared kinases and phosphatases between metabolites and phosphosites lead to a slight increase of the ratios. However, at the same time the number of outliers increases, too. (Figures 11 (a) to (d)). Therefore, since an analysis with a minimum requirement of five shared kinases and phosphatases best separates from random with respect to the significant data p-values $\leqslant 0.05$, this number seems to be a reasonable choice for future analyses.

Figures 12 to 18 illustrate examples of ion–phosphosite pairs showing significant negative or positive correlations between their abundance fold changes. For each data p-value threshold (ranging from 0.01 to 0.4 and incremented by steps of 0.025) both Pearson correlation and Spearman rank correlation were calculated.

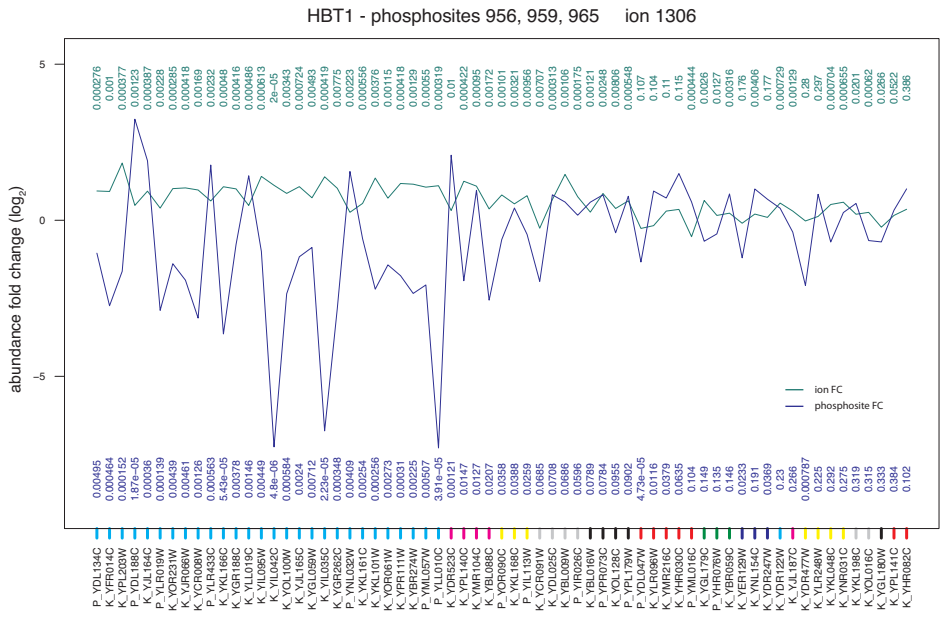
Figure 12 features ion '1306', which could be mapped to *UDP–N–acetyl–D–glucosamine* (KEGG [130] compound C00043 [131]), a compound involved in amino sugar and nucleotide sugar metabolism (KEGG pathway sce00520 [132]). The abundance changes of this molecule clearly (anti–)correlate with the abundance changes of a triple–serine–phosphorylated peptide (amino acids 956, 959, and 965) of Hbt1, a protein known to be a substrate of the Hub1p ubiquitin–like protein ([133], and see Figure 12(a), and Table 12(b) for correlation values).

Of the three correlated residues of Hbt1, only phosphosites 956 and 959 are currently curated in PhosphoGRID [15], but neither of them is functionally annotated. Equally little is known about the function of Hbt1 itself.

However, our correlation analysis links the protein to the metabolic network and suggests a functional relationship between Hbt1 and *UDP–N–acetyl–D–glucosamine*. This observation is supported by the STRING database [94]. As shown in figure 12(c), STRING indicates a direct PPI between Hbt1 and Mcr1, as well as between Hbt1 and Hxk1, two enzymes (*mitochondrial NADH–Cytochrome b5 reductase*, EC 1.6.2.2, and *hexokinase isoenzyme 1*, EC 2.7.1.1, respectively) known to participate in our pathway of interest ([132]).

Figure 13 depicts another example for an anti–correlation observed between the abundance fold changes of a metabolite and a phosphopeptide.

(a)



(b)

data p-value cutoff	Pearson correlation	Pearson p-value	Spearman correlation	Spearman p-value
0.01	-0.4926	0.007741	-0.4773	0.01098
0.025	-0.4669	0.007057	-0.415	0.01894
0.05	-0.4842	0.003204	-0.481	0.003812
0.075	-0.3671	0.02151	-0.4002	0.01211
0.1	-0.391	0.009536	-0.4477	0.002854
0.125	-0.4098	0.003821	-0.4893	0.0004899
0.15	-0.4196	0.002176	-0.5032	0.0002041
0.2	-0.4232	0.001432	-0.494	0.0001784
0.25	-0.4246	0.001233	-0.4955	0.0001463
0.275	-0.4251	0.001091	-0.4948	0.0001298
0.3	-0.4106	0.001118	-0.4615	0.0002418
0.325	-0.415	0.0007965	-0.4547	0.000239
0.35	-0.406	0.000962	-0.4423	0.0003253
0.4	-0.4144	0.0005987	-0.4429	0.0002541

(c)

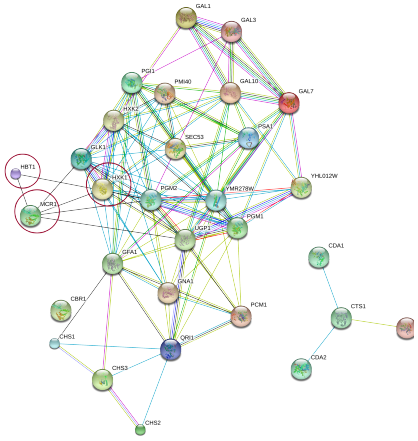


Figure 12: *Correlation analysis between phosphosites 956, 959, 965 of Hbt1 and UDP–N–acetyl–D–glucosamine.* **(a)** Fold change plot of metabolite and phosphopeptide abundances. For each shared kinase/phosphatase (x–axis) the corresponding fold changes ($\log_2$, y-axis) in ion abundance (green) and in phosphopeptide abundance (blue) are plotted. Kinases/phosphatases tagged in the same color were considered together for calculating the fold change correlations at a given data p-value cutoff. **(b)** Correlation values relative to different data p-values. At a given data p-value, all fold changes with a less or equal data p-value were take into account for the correlation calculations. **(c)** STRING network showing direct interactions between Hbt1 and metabolic enzymes Mcr1 and Hxk1.

Here, the ion involved is annotated to the metabolite *uracil* (KEGG compound C00106 [134]). In contrast, the correlated di–phosphorylated peptide (serine 222 and threonine 225) belongs to a yet uncharacterized protein, YHR097C. The strong and significant anti–correlation found (see Table 13(b) for correlation coefficients and p-values) indicates a connection of YHR097C to metabolism. Again, the STRING database supports this assumption and states an interaction with both *aldehyde dehydrogenase* (Ald4, EC 1.2.1.3) and *glutamate decarboxylase* (Gad1, EC 4.1.1.15) as depicted in Figure 13(c). Remarkably, the two enzymes are participating in β–alanine metabolism (KEGG pathway sce00410 [135]), the very pathway which synthesizes *uracil*, and thus encourage our hypothesis of an association of YHR097C to metabolism.

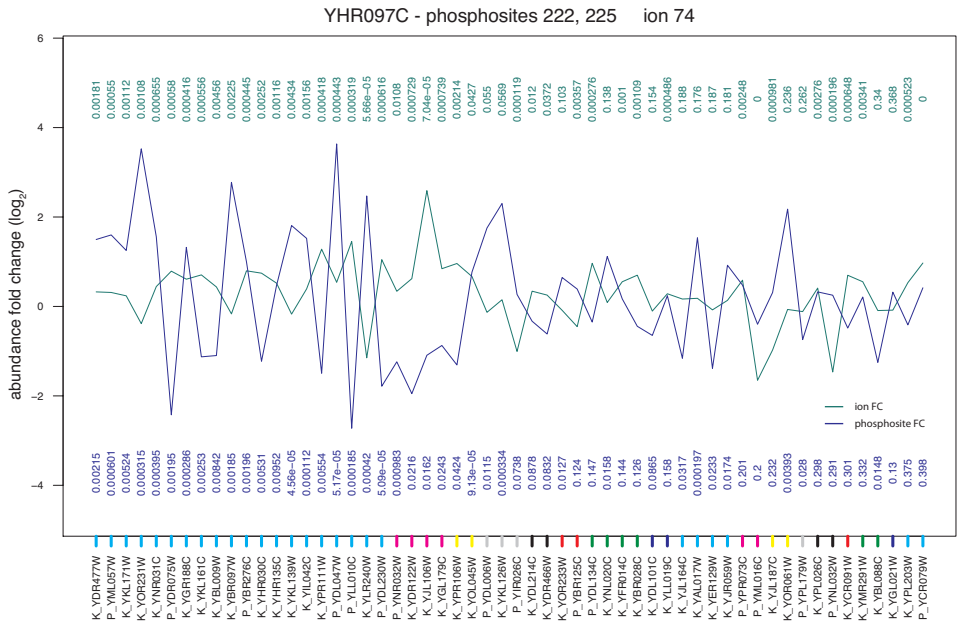
The impact of TFs on metabolomic regulation triggered by protein phosphorylation is exemplified in Figure 14. Here, a highly significant correlation between a double–phosphorylated peptide (residues 322 and 326) of Not3, a subunit of the Ccr4–Not TF complex [136], and ion '141', a small molecule annotated as *(S)–Malate* (KEGG compound C00149 [137]) and participating in methane metabolism (KEGG pathway sce00680 [138]), is illustrated (Figure 14(a)). Though no direct interactions with enzymes involved in methane metabolism are known, the STRING database connects Not3 to the methane pathway protein network via the enzyme Pdc1 (*pyruvate decarboxylase*, EC 4.1.1.1, see Figure 14(c)), confirming the relationship between Not3 and methane metabolism identified in our correlation analysis.

As mentioned in the beginning of this section, most of the ions quantified have not been annotated yet. To bypass this limitation, we employed STITCH, a database for known interactions between proteins and chemical compounds [139]. Using the information provided by STITCH, we were able to locate and map some of the unidentified ions to known metabolites based on their m/z values.

In doing so, ion '205' (m/z 157.051) could be assigned to *7,8–dihydropteroate* (KEGG compound C00921 [140]), a small molecule synthesized during folate biosynthesis (KEGG pathway sce00790 [141]). The corresponding reaction is catalyzed by the enzyme *dihydropteroate synthase* (Fol1, EC 2.5.1.15). In our analysis, we found the two molecules indeed to be interrelated: the metabolite's abundance fold changes show a significant anti–correlation to the fold changes of tyrosine residue 33 of the protein (see Figure 15), strongly supporting our annotation of the metabolite.

Another example of mapping unidentified ions to pathways is shown in Figures 16 to 18. Using STITCH, ion '1612' (Figure 16) was identified as metabolite *acetyl–CoA* (KEGG compound C00024 [142]), ion '1559' as *acetoacetyl–CoA* (KEGG compound C00332 [143], Figure 17), and ion '1510' as *coenzyme A* (KEGG compound C00010 [144], Figure 18). The three molecules are participating in neighboring and interlinked metabolic pathways,

(a)



(b)

data p-value cutoff	Pearson correlation	Pearson p-value	Spearman correlation	Spearman p-value
0.01	-0.7337	0.0002315	-0.803	2.385e-05
0.025	-0.6418	0.0007229	-0.7235	9.975e-05
0.05	-0.6454	0.0003708	-0.7395	2.637e-05
0.075	-0.6005	0.0005731	-0.7276	1.392e-05
0.1	-0.5865	0.0005252	-0.7093	1.411e-05
0.125	-0.5708	0.0005229	-0.6581	4.817e-05
0.15	-0.5789	0.0001743	-0.6769	7.682e-06
0.175	-0.5577	0.0002252	-0.6457	1.467e-05
0.2	-0.5159	0.0003987	-0.5657	0.0001022
0.225	-0.431	0.003121	-0.5267	0.0002489
0.25	-0.4222	0.003118	-0.5165	0.000247
0.275	-0.4098	0.003822	-0.4967	0.000391
0.3	-0.3836	0.005955	-0.4728	0.0006059
0.325	-0.3873	0.004991	-0.4838	0.0003794
0.35	-0.3718	0.006117	-0.4556	0.0006917
0.375	-0.3715	0.005676	-0.4475	0.0007815
0.4	-0.3675	0.005325	-0.4314	0.001003

(c)

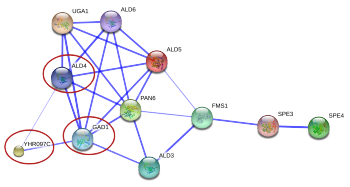


Figure 13: *Correlation analysis between phosphosites 222, 225 of uncharacterized protein YHR097C and uracil* **(a)** Fold change plot of metabolite and phosphopeptide abundances. For details, see 12(a). **(b)** Corresponding correlation values. **(c)** STRING network showing direct interactions of YHR097C with Ald4 and Gad1, two metabolic enzymes that participate in β–alanine metabolism, just like the metabolite of interest, *uracil*, does.

among others in the fatty acid biosynthesis (KEGG pathway sce00061 [145]) and fatty acid metabolism (KEGG pathway sce00071 [146]). The multifunctional enzyme *acetyl–CoA carboxylase* (Acc1, EC 6.4.1.2) plays a prominent role in these pathways and is the key enzyme of fatty acid biosynthesis by catalyzing the carboxylation of *acetyl–CoA* in order to produce *malonyl–CoA* [147].

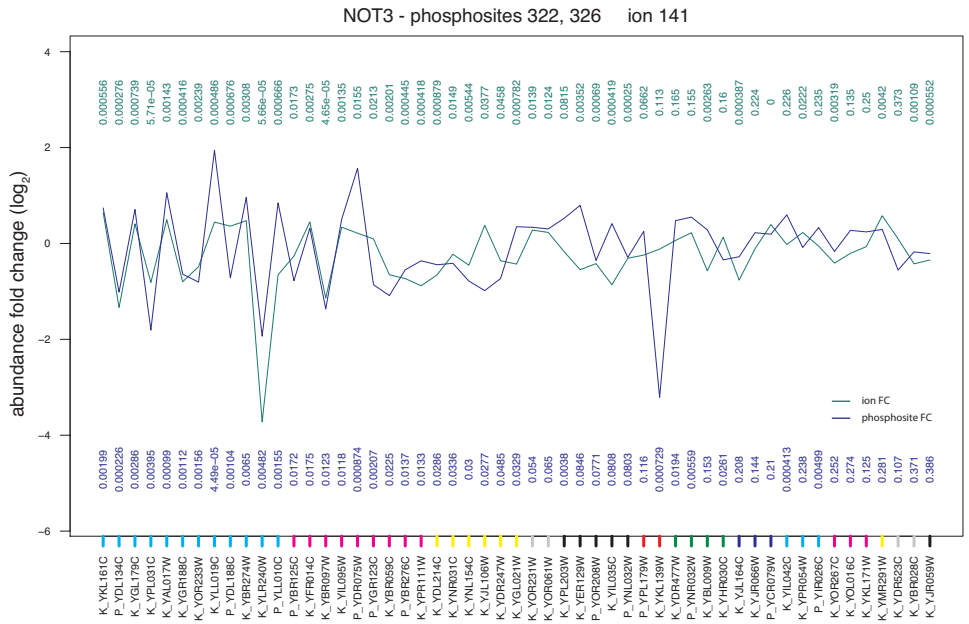
The functional relationship between Acc1 and the three compounds is also visible in our analysis. All three metabolites show a sufficient correlation in their abundance fold changes to serine residue 1159 of Acc1. *Acetyl–CoA* and *coenzyme A* are positively correlated, whereas *acetoacetyl–CoA* is negatively correlated to serine 1159. It should be noted, however, that the correlations are only significant when calculated based on extended data p-values (for correlation values see Tables 16(b), 17(b), and 18(b)).

Some details of the regulation of Acc1 are known. For instance, the enzyme is deactivated through phosphorylation by the AMP–activated kinase Snf1. In contrast, the dephosphorylation by phosphatase Sit4 leads to the activation of Acc1 [148, 149].

Since Acc1 was not detected as a target of Snf1 in the phosphoproteome dataset, a likely association between *acetyl–CoA*, the substrate of Acc1, and Snf1 could not be verified in our correlation analysis. Nevertheless, the positive correlation between Acc1 and *acetyl–CoA* reflects the molecular background, as Acc1 is only active upon dephosphorylation. Interestingly, Sit4 (YDL047W), the phosphatase which is responsible for the activation of Acc1, is among the perturbations that were considered for the correlation analysis. Its knockout yields a positive fold change of the abundance of serine 1159 of Acc1, an outcome that is consistent with the fact that Acc1 constitutes as a direct target of Sit4.

Acetoacetyl–CoA is synthesized from *acetyl–CoA* in a reversible reaction. During this process, *coenzyme A* is released as well, and acts as an inhibitor of the reaction ([150, 151]). The negative correlation of *acetoacetyl–CoA* with Acc1 could be explained in light of the relations between the metabolites themselves. On the other hand, Acc1 is a multifunctional enzyme and also involved in other pathways, as are the various metabolites. These factors as well as the still not fully understood complex regulation of Acc1 [148] impede more definitive conclusions.

(a)



(b)

data p-value cutoff	Pearson correlation	Pearson p-value	Spearman correlation	Spearman p-value
0.01	0.731	0.006908	0.8042	0.002746
0.025	0.6988	0.000425	0.7221	0.0003251
0.05	0.6433	0.0002951	0.6374	0.0004693
0.075	0.6523	0.0001259	0.6488	0.0001985
0.1	0.6089	0.0001329	0.5575	0.0007504
0.125	0.5062	0.001633	0.5125	0.001609
0.175	0.5054	0.0008792	0.5266	0.0005786
0.225	0.5076	0.0005113	0.5211	0.0004107
0.25	0.509	0.0003031	0.5213	0.0002499
0.275	0.5098	0.0001832	0.5199	0.0001605
0.3	0.512	0.0001445	0.5189	0.000141
0.375	0.5047	0.000136	0.5038	0.0001719
0.4	0.5049	0.0001157	0.504	0.0001467

(c)

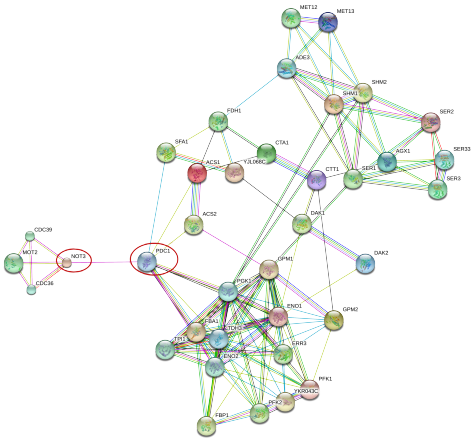
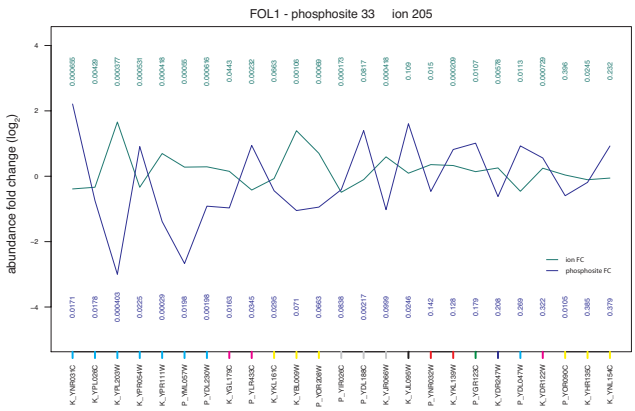


Figure 14: *Correlation analysis between the TF Not3 and (S)–Malate* **(a)** Fold change plot of metabolite and phosphopeptide abundances. For details, see 12(a). **(b)** Corresponding correlation values. **(c)** STRING network showing an indirect interactions of Not3 to the methane pathway protein network via enzyme Pdc1.

(a)

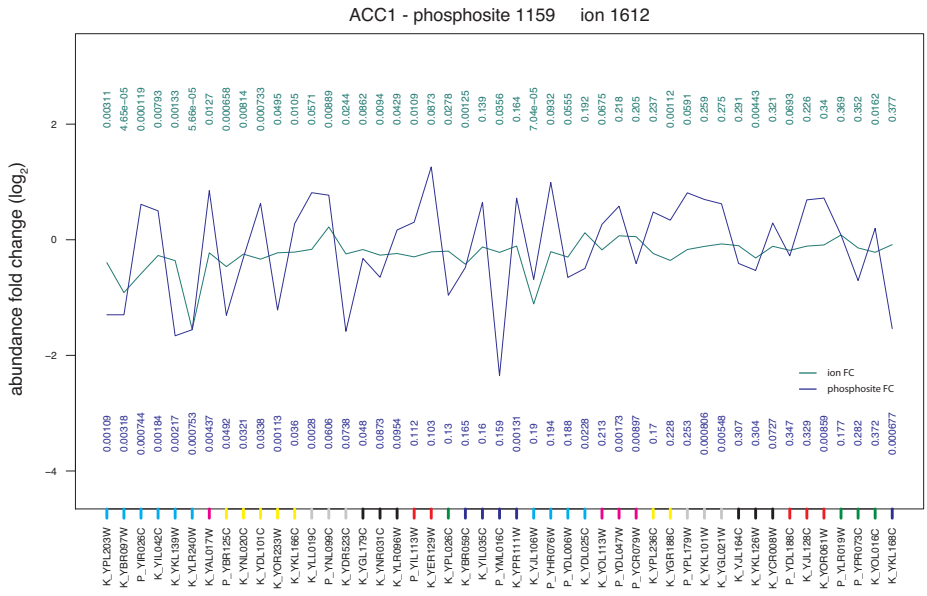


(b)

data p-value cutoff	Pearson correlation	Pearson p-value	Spearman correlation	Spearman p-value
0.025	-0.7742	0.04105	-0.8571	0.02381
0.05	-0.7962	0.01024	-0.8833	0.003075
0.075	-0.7177	0.008583	-0.8392	0.001192
0.1	-0.694	0.0041	-0.8179	0.0002971
0.125	-0.6659	0.004862	-0.7824	0.0005334
0.15	-0.6438	0.003934	-0.7276	0.0009053
0.2	-0.6363	0.003399	-0.7158	0.0008175
0.225	-0.6356	0.002595	-0.7218	0.0004803
0.275	-0.6528	0.001336	-0.7195	0.0003496
0.325	-0.6457	0.001172	-0.7143	0.0002788
0.4	-0.6386	0.0005918	-0.6815	0.0002537

Figure 15: *Correlation analysis between phosphosite residue 33 of Fol1 and ion '205', annotated as 7,8–dihydropteroate using STITCH* **(a)** Fold change plot of metabolite and phosphopeptide abundances. For details, see 12(a). **(b)** Corresponding correlation values.

(a)

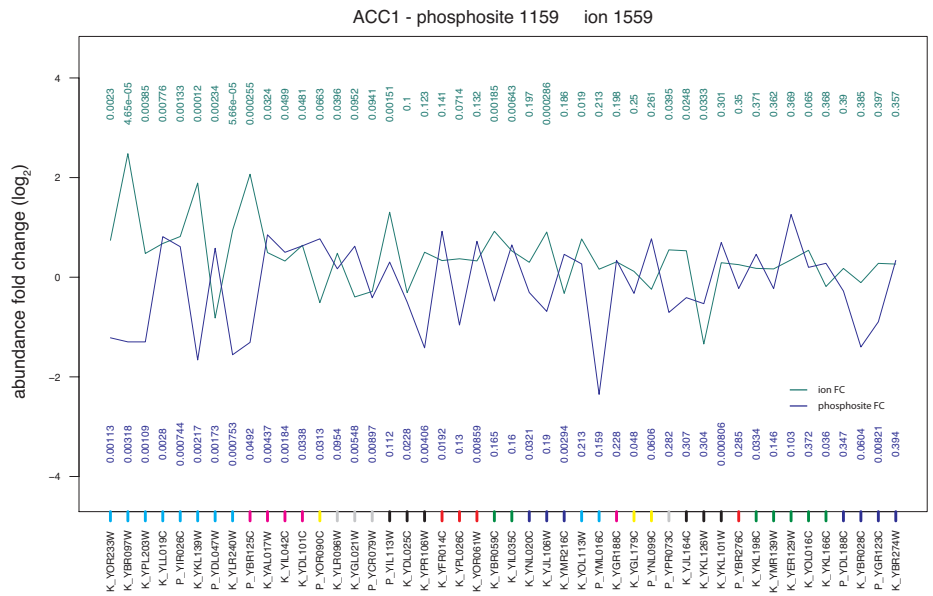


(b)

data p-value cutoff	Pearson correlation	Pearson p-value	Spearman correlation	Spearman p-value
0.01	0.4022	0.4293	0.1449	0.7841
0.025	0.5113	0.2408	0.4685	0.289
0.05	0.4493	0.1428	0.4799	0.1144
0.075	0.4862	0.06609	0.5344	0.04014
0.1	0.4815	0.04304	0.5142	0.02903
0.125	0.4834	0.03081	0.5258	0.01727
0.15	0.4585	0.03659	0.48	0.02764
0.175	0.4135	0.03993	0.4947	0.01194
0.2	0.3833	0.04014	0.4858	0.007551
0.225	0.3945	0.02547	0.4696	0.006692
0.25	0.3902	0.02252	0.4469	0.00805
0.275	0.4131	0.01106	0.496	0.001797
0.325	0.41	0.008596	0.4689	0.002275
0.35	0.4221	0.004811	0.4897	0.0008582
0.375	0.4121	0.004428	0.4436	0.002015
0.4	0.3784	0.008732	0.3854	0.007465

Figure 16: *Correlation analysis between residue 1159 of Acc1 and ion '1612', annotated as acetyl–CoA using STITCH* **(a)** Fold change plot of metabolite and phosphopeptide abundances. For details, see 12(a). **(b)** Corresponding correlation values.

(a)

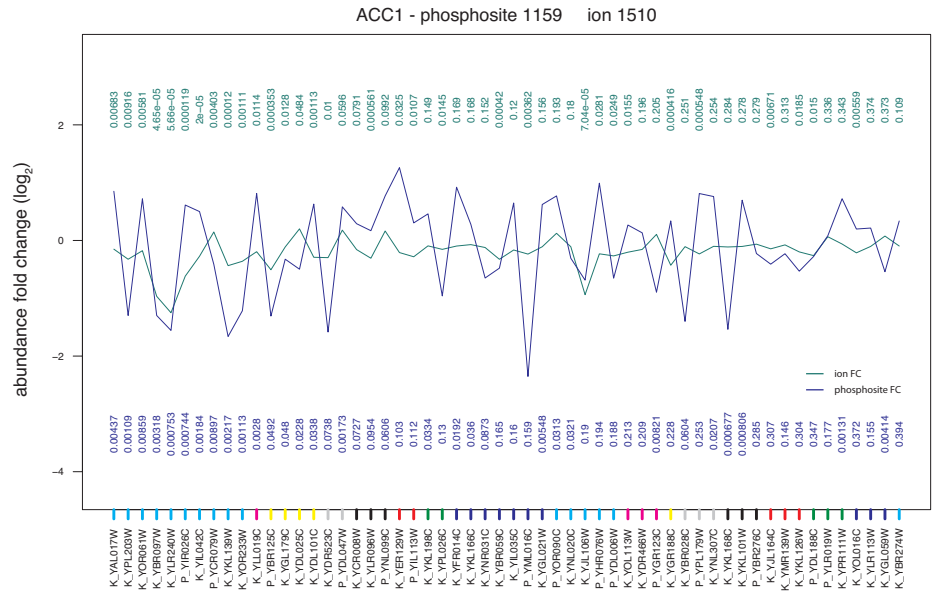


(b)

data p-value cutoff	Pearson correlation	Pearson p-value	Spearman correlation	Spearman p-value
0.01	-0.5713	0.1391	-0.5509	0.157
0.05	-0.6254	0.02965	-0.5499	0.06398
0.075	-0.6649	0.01315	-0.575	0.03982
0.1	-0.6466	0.006791	-0.518	0.03983
0.125	-0.5325	0.01892	-0.4581	0.04856
0.15	-0.5131	0.01461	-0.4637	0.02973
0.175	-0.5095	0.01098	-0.4497	0.02748
0.2	-0.5237	0.005056	-0.4546	0.01721
0.225	-0.4221	0.02257	-0.3577	0.05678
0.25	-0.4253	0.01915	-0.3549	0.05432
0.275	-0.4406	0.01161	-0.3621	0.04172
0.3	-0.4396	0.01047	-0.3714	0.03334
0.325	-0.3751	0.02417	-0.3394	0.04284
0.35	-0.3745	0.02241	-0.3252	0.04954
0.375	-0.3718	0.01534	-0.2965	0.05655
0.4	-0.3283	0.02591	-0.2247	0.1333

Figure 17: *Correlation analysis between residue 1159 of Acc1 and ion '1559', annotated as acetoacetyl–CoA using STITCH* **(a)** Fold change plot of metabolite and phosphopeptide abundances. For details, see 12(a). **(b)** Corresponding correlation values.

(a)



(b)

data p-value cutoff	Pearson correlation	Pearson p-value	Spearman correlation	Spearman p-value
0.01	0.4821	0.1583	0.6261	0.05278
0.025	0.5095	0.1094	0.6651	0.02553
0.05	0.4436	0.09766	0.5416	0.03706
0.075	0.4581	0.06443	0.5138	0.03488
0.1	0.5025	0.02395	0.5182	0.01924
0.125	0.486	0.02184	0.5061	0.01626
0.15	0.472	0.01989	0.4658	0.02181
0.175	0.4353	0.01437	0.4375	0.01383
0.2	0.4446	0.006594	0.4648	0.00429
0.225	0.4093	0.009662	0.4103	0.009485
0.25	0.3974	0.01111	0.3912	0.01255
0.275	0.3729	0.01377	0.349	0.0218
0.3	0.3499	0.01712	0.316	0.03243
0.325	0.3437	0.0156	0.2974	0.03799
0.35	0.352	0.01049	0.3058	0.02746
0.375	0.3392	0.0113	0.2768	0.04074
0.4	0.3418	0.009923	0.2758	0.03963

Figure 18: *Correlation analysis between residue 1159 of Acc1 and ion '1510', annotated as coenzyme A using STITCH* **(a)** Fold change plot of metabolite and phosphopeptide abundances. For details, see 12(a). **(b)** Corresponding correlation values.

11.3 CONCLUSION AND PERSPECTIVES

Our preliminary results illustrate the potential of an integrated analysis of perturbed metabolome and phosphoproteome data in order to elucidate the role of protein phosphorylation on metabolome regulation.

The examples presented show the usefulness of our approach in several aspects. First, the correlation analysis can contribute significantly to the functional annotation of proteins. So far, both Hbt1 and YHR097C had not been linked to metabolic pathways before. The strong correlations to metabolites we could identify in our analysis suggest such metabolic associations, which are indeed supported by findings and predictions in the STRING database. Secondly, besides protein annotation, our analysis can also be employed for ion annotation. Combined with a database providing context information such as STITCH, it is possible to locate the ion based on its m/z value and its relationship (in our case correlation) to a protein. Thirdly, our analysis can reveal phosphosites that act as indicator sites of protein function, like enzyme activity or transcriptional control activity.

It is worth mentioning that our analysis was implemented in an unbiased fashion, meaning that we considered the kinase and phosphatase knockouts as anonymous perturbations and did not employ knowledge about known events of kinase/phosphatase regulation. Despite this, the results appear reasonable and meaningful, further validating our analysis.

Besides applications already mentioned, our correlation analysis could also be helpful for the identification of functional phosphosites that were unidentified so far. For example, a strong correlation to an ion may indicate a functional role of a phosphosite whether or not it was reliably quantified initially. Consequently, our analysis could act as a filter in order to detect functional phosphosites previously unknown to be connected to a given kinase or phosphatase.

Part IV

APPENDIX

A

AN IN VIVO GENOME-WIDE STUDY OF ORGAN GROWTH REGULATION

The following preliminary manuscript describes a genome–wide *in vivo* RNA interference (RNAi) screen in order to identify new growth regulators in the fruitfly *Drosophila melanogaster*. I implemented the full bioinformatic pipeline in order to analyze the HTP data. Furthermore, I compiled and prepared data from a number of other RNAi screens in order to compare their results to our data.

The provisional manuscript is attached below.

An *in vivo* Genome-Wide Study of Organ Growth Regulation

Gerlinde Reim, Stefanie Wanka, Christian von Mering and Konrad Basler

Institute of Molecular Life Sciences

University of Zurich

Winterthurerstrasse 190

CH-8057 Zurich

Switzerland

Contact: Konrad Basler

phone: +41-44-635-3111

fax: +41-44-635-6864

email: basler@imls.uzh.ch

Key Words

Growth, organ size, *Drosophila melanogaster,* wing imaginal disc, Hedgehog, Decapentaplegic, RNAi

Summary

A major problem concerning organ development is how tissue growth is achieved, and how its regulation is coordinated with other biological processes such as cell growth, cell division, and differentiation. Various high-throughput RNAi screens focusing on cell growth and proliferation have relied on cultured cells. In order to learn more about organ growth we used the wing of *Drosophila melanogaster* as a model system to uncover new growth regulators. We could thus study growth control in an entire, intact organ. Hence we conducted a genome-wide *in vivo* RNAi screen in the wing system with a coverage of 11'250 genes (83% of *Drosophila* protein coding genes). 21'014 transgenic RNAi lines were assayed in two differently sensitized tester backgrounds exhibiting small or large wings due to insufficient or excessive levels of Hh signaling, respectively. The Hedgehog (Hh)- and Decapentaplegic (Dpp)/TGFβ-signaling pathways are central to growth control in wing imaginal discs but it is unclear how they impinge on growth and if/how they define disc size. We have identified known as well as novel growth regulators. 2'669 genes (23% of all tested genes, 19.6% of all *Drosophila* protein-coding genes) revealed a wing growth phenotype upon RNAi in at least one of the two conditions tested, including genes with an established function in wing disc growth. Our *in vivo* screen therefore provides a comprehensive catalogue of growth regulatory genes, revealing many novel components of growth that have not been so far recognized in cell culture screens. A large number of our candidate genes have mammalian orthologs. We show that *CG6854* acts as a positive growth regulator during wing disc development while *CG14542* are required to restrict growth, illustrating the potential of this screen to uncover novel growth regulators important for the control of cell proliferation and organ size in humans.

Introduction

Tissue growth and its orchestration with morphogenetic processes are central to the development of multicellular organisms. The size of an organ or appendage is determined by the number and size of the cells it contains (Reis and Edgar, 2004; Cook and Tyers, 2007). Pro-growth processes such as cell proliferation (cell division) are counteracted by specific inhibitors, or by cell death. The final size of an organ is determined by extracellular and intracellular mechanisms that induce and terminate growth. These mechanisms remain largely elusive as concurrent with growth, different gene expression programs need to be implemented to ensure proper patterning and differentiation of the organ.

The ideal system to study growth regulation is a tissue that contains mitotically active and differentiating cells amenable to genetic manipulation *in vivo*, but which is largely dispensable for the viability of the organism. *Drosophila* wing imaginal discs, anlagen of the adult wings, satisfy these criteria. During the first larval day a 50-cell wing primordium starts to proliferate and increases thousand-fold in size by the end of larval development (five days later), prior to undergoing metamorphosis into adult wing cuticle. This growth is influenced (i) by extrinsic parameters, including nutrients, oxygen, temperature and hormone signaling (Britten and Edgar, 1998; Kawamura et al., 1999; Leopold and Perrimon, 2007), and (ii) by organ-autonomous, intrinsic factors (Bryant and Simpson, 1984; Neto-Silva et al., 2009). Morphogen-operated pathways that regulate wing patterning are emerging as key players in the intrinsic control of growth. For example, the Hh signaling cascade is crucial for patterning the A/P (antero-posterior) wing axis and for controlling wing growth (Tabata and Takei, 2004; Crozatier, 2004). The key growth effector of Hh signaling is *dpp*, whose expression is activated at the A/P border (Affolter and Basler, 2007). Dpp exerts its pro-growth activity, primarily by repressing the growth-antagonist *brinker* (Campbell and Tomlinson, 1999; Jazwinska et al., 1999; Minami et al., 1999; Marty et al., 2000; Cook et al., 2004; Martin et al., 2004; Schwank et al., 2008). Disrupted Hh signaling causes many diseases ranging from developmental abnormalities to an increasing list of cancers (Wetmore, 2003; Rubin and de Sauvage, 2006). In addition to morphogens, mechanical forces in the disc may also affect growth processes (Hufnagel et al., 2007; Aegerter-Wilmsen et al., 2007). Due to their genetic amenability imaginal discs have served to uncover many aspects of growth cues, and their networks form the basis of our still incomplete understanding of growth control at the organ level.

Traditional forward genetic methods, such as mutagen-based loss-of-function screens, have revealed a number of core components of growth regulatory pathways. More recently, high-throughput screening approaches using RNAi in cell culture-based assays have provided further exciting insights (DRSC http://flyrnai.org; DasGupta and Gonsalves, 2008; Mathey-Prevot and Perrimon, 2007; Fuchs and Boutros, 2006; Lents and Baldassare, 2006). However, a more powerful approach would be to investigate the consequences of loss of function of growth-relevant genes in the context of the entire animal or whole organ.

One way to accomplish this are reverse-genetic approaches like “gene-knockouts” by homologous recombination (Rong et al., 2002); however due to their laboriousness these techniques have not been

employed for large-scale studies. Moreover, global loss of gene function often causes lethality of homozygous mutant animals. While this difficulty can be circumvented by mosaic analyses (Garcia-Bellido and Dapena, 1974, Tyler et al., 2006) this approach is extremely time-consuming.

An alternative is RNA interference that has become the standard methodology to investigate the loss or reduction of gene function in a high-throughput manner. Numerous screens were performed in cell culture with the aim to catalog genes involved in diverse processes, ranging from cell cycle control to lipid droplet formation (Boutros et al., 2004; Björklund et al., 2006; Guo et al., 2008; Steinbrink and Boutros, 2008). The generation of a genome-wide transgenic RNAi library by the VDRC (Vienna Drosophila RNAi Center; Dietzl et al., 2007) enabled the performance of such screens in a whole organ or organism. By using this library we achieved a systematic and unbiased *in vivo* loss-of-function screen to uncover growth regulators of an entire, intact organ - the *Drosophila* wing. *Drosophila's* low gene redundancy further simplifies a functional screening approach to identify novel regulators of growth. Advantageous to previous cell based approaches, performing a screen in an intact organ should reveal more components relevant for tissue growth. We could assay 11'250 genes (83% of all *Drosophila* protein coding genes), which are represented by 21'014 transgenic "UAS-hairpin" lines (1.8 lines per gene on average). 2'669 genes (23% of all tested genes, 19.6% of *Drosophila* protein-coding genes) revealed a wing size phenotype upon RNAi, including known genes with an established function in wing disc growth. In addition to wing size changes, phenotypes relating to vein patterning and epithelial morphology could be distinguished. Our approach is accompanied by a careful bioinformatic analysis of the data. Genes that scored positively in our screen have a significantly enriched set of human orthologs, validating the use of the *Drosophila* wing epithelium, which resembles epithelia of higher metazoans, as a system to discover regulators important in tumorigenesis and mammalian organs growth.

Results and Discussion

Establishment of two sensitized systems to identify growth regulators *in vivo*

Successful *in vivo* screening approaches require a sensitive, robust, and easily identifiable phenotype. To achieve this and to maximize the recovery of growth regulators we created two sensitized tester systems with larger or smaller wings by increasing or decreasing, respectively, the levels of Hh signaling. We employed two transgenes expressing different forms of Cubitus interruptus (Ci): Ci^{rep} or Ci^{act}. Ci is the transcriptional effector of Hh signaling (Slusarski et al., 1995; reviewed in Huangfu and Anderson, 2006) and is proteolytically truncated in cells that do not receive the Hh signal. The full length form is favored upon Hh signaling and is a transcriptional activator. Importantly, *dpp* is activated by Hh signaling and repressed in absence of Hh. Dpp signaling in turn largely accounts for wing growth (Affolter and Basler, 2007). In our screen system the Ci^{rep} transgene encodes a constitutively truncated repressor (Methot and Basler, 1999), while Ci^{act} lacks phosphorylation sites necessary for the proteolytic cleavage and is thus always in the active, full length form irrespective of the presence of Hh (Methot and Basler, 2000). Consequently, these transgenes caused either small wings (*nub-Gal4, UAS-ci^{rep}*), or large wings (*nub-Gal4, UAS-ci^{act}*) (Fig. 1; Fig. 2B-B'; Material and Methods). Since Hh/Dpp signaling acts predominantly along the A/P axis the Ci^{act} tester wings are significantly enlarged along the A/P axis, and less pronounced along the PD (proximo-distal) axis. In addition to being suited for detecting suppressors of their respective phenotypes (i.e. enlarging the small wings or reducing wing over-size), both wing phenotypes could also be enhanced. Therefore in each of the setups it is possible to identify both positive and negative regulators of growth. This system is hence not only able to identify potent regulators of wing size, but should also reveal more subtle effectors. This is particularly important in cases of functional redundancy or moderate RNAi-mediated knockdown causing hypomorphic phenotypes. The phenotypes of the tester lines were robust and showed only limited variability.

We expressed the VDRC RNAi transgenes in the primordial wing pouch by *nubbin*-Gal4 (*nub*-Gal4; Fig. 1). The pouch of the wing imaginal disc develops into the wing proper. As a consequence, effects on primordial wing growth could be observed in the adult wing. The *nub*-Gal4 driver is active from the 2nd larval stage onwards in the wing pouch, during the exponential growth phase until early pupal stages (Calleja et al., 1996; Ng et al., 1996; Cifuentes and Garcia-Bellido, 1997).

Validation of the screening setup

Our tester system is based on modulating the levels of Hh signaling. Ci is one of the most downstream components in this cascade, and the transcriptional output of Ci^{rep} or Ci^{act} should be independent of the Hh signaling input. Nevertheless, the endogenous Hh-regulated Ci may contribute to the readout – an ideal situation since our screen could potentially uncover also novel Hh signaling components. This was indeed the case since depletion of positive Hh signaling components that act functionally upstream of Ci, like Smoothened (Smo) or Hh itself, caused a size reduction of Ci^{rep} and Ci^{act} wings predominantly along the

A/P axis due to the pronounced activity of Hh/Dpp signaling along this axis (Fig. 2C,C'; not shown). A ci^{RNAi} hairpin, which targets both transgenes as well as the endogenous transcript, caused size reduction of large tester wings (Fig. 2D), however, no further size reduction in the small tester wings (Fig. 2D'). This was not unexpected since the size increase due to the removal of transgenic Ci^{rep} would be counteracted by the loss of endogenous Ci activator. To further validate the screening setup we tested the effect of knocking down a number of genes known to act in signaling pathways other than the Hh pathway that are important during wing disc development. Dpp is a key regulator of wing growth, and in our tester conditions *dpp* expression levels were up- or downregulated upon Ci^{act} and Ci^{rep}, respectively. As expected, RNAi mediated depletion of transcripts encoding Dpp signaling components such as *mad*, *med*, *put, tkv* or *brk*, modified the ci^{act} and ci^{rep} phenotypes in agreement with their sign of action in the Dpp pathway (Fig. S1 C-G'). In sum, of the well-established components of Hh (27 genes) and Dpp (15 genes) signaling addressed by RNAi we found 16 and 14 genes, respectively, with the expected phenotype (Fig. 2, Fig. S2, Tab. S1). Among the 24 Hh signaling pathway genes, 4 were lethal before eclosure, 1 resulted in an unexpected phenotype, and 3 showed no phenotype upon RNAi, including 2 genes encoding Hh secretory components (*ski*, *disp*), where a strong phenotype would not be expected in the Ci^{act} condition. Based on RNAi phenotypes obtained from a control gene set (Fig. S1) we estimate that 81.4% of genes show the expected phenotype upon RNAi, whereas 11.6% are false negative. We found 7 % false positive genes which is likely based on off-target effects and/or mutagenic effects caused by a hairpin-transgene insertion.

We also tested Hh-independent growth regulatory pathways, such as components of Insulin- and Hippo-signaling. Depletion of *pten*, a negative regulator of Insulin signaling, caused enlarged wings in both tester conditions without any effect on vein patterning (Fig. 2E,E'). Similarly, targeting *merlin, expanded* and *warts* also caused increased wing sizes in both tester conditions (Fig. 2F,F'; Fig. S1H-I'); These components normally inhibit Hippo signaling, which has emerged as an important repressor of cell proliferation and promoter of apoptosis in differentiating epithelial cells (reviewed in Saucedo and Edgar, 2007; Reddy and Irvine, 2008). In summary, our assay system has the potential to uncover, by RNAi mediated depletion, both positive and negative regulators of growth.

Screen results: the yield of the wing size screen

We screened a total of 11'606 genes (21'014 lines), which represent a coverage of approximately 83% of all predicted 14'306 protein-coding genes in *Drosophila*. RNAi of 3'677 genes (5067 RNAi lines) caused alterations of at least one tester phenotype (Fig. 3A). We eliminated 73 precarious genes from our analysis (Suppl. Info). We categorized our positive gene set (3'677 genes) into different phenotypic classes (Fig. 3B), whereas Tab. S2 details the genes associated with distinct phenotypes and distinguishes between the two different tester conditions. The hierarchy used for phenotype categorization is described in the Suppl. Info.

We scored 2'596 genes to affect wing size in at least one tester condition (Fig. 3B; Tab. S2). The main class consists of growth activators (1'728 genes), which reduced tester wing sizes upon RNAi mostly in a proportional (allometric) manner (Fig. 4D-D''). RNAi of 412 genes resulted in a round wing phenotype (Fig. 4E,F). Roundish wings can either result from growth defects predominantly along the wing proximo-distal axis (Fig. 4F), or from aberrant cell orientation as observed in mutants with misoriented cell divisions like *dachsous* (Strutt, 2005), and even in the context of overgrowth, as seen in strong Hh/Dpp expressing wings (Fig. 4E). From the plain, roundish wings we distinguished an other roundish class for its additional dusky and convex-bent wing epithelia based on altered wing hair densities (Fig. 4F'). A last apparently small-wing class characterized by lack of wing tissue comprises 86 genes which showed wings marginal notches to different extents (Fig. 4D'''). In addition to small-wing phenotypes we also found growth repressors (370 genes), as revealed by wing enlargement upon RNAi (Fig. 4E,E').

In addition to growth phenotypes we also observed altered wing morphologies, consisting of three main classes: (i) wings with blistered or wrinkled cuticle (115 genes; Figure 4G), (ii) wings with a singed appearance (72 genes; Fig. 4G'), and (iii) wings which did not properly unfold (25 genes; Fig. 4G'''; Suppl. Info). Blistered or wrinkled cuticles were frequently associated with a size phenotype (Fig. 4D''). Furthermore, venation- (Fig. 4H) and wing hair defects (Fig. 4H'), including supernumerary or aberrantly spaced trichomes, were also mostly associated with size-phenotypes. This is concurrent with the acitivities of several signaling pathways during wing development to provide not only growth-, but also patterning cues.

Finally, we observed lethality phenotypes upon RNAi, including necrotic or apoptotic wings (88 genes), and lethality at pupal or larval stages (909 genes; Suppl. Info). One cause for the small size of the Ci^{rep} tester wing is apoptosis, as Caspase-3 is hyperactivated in such wing primordia (Fig.S 2B,B', compare to A,A'). Ci^{rep} tester were therefore sensitive to detect pro-apoptotic genes, whereas Ci^{act} wings remained unaltered since they do not display apoptosis. Unlike several cell-culture growth screens we identified well established pro-apoptotic genes: RNAi of *hid* and *rpr* led to a gain of Ci^{rep} wing size, and a decrease of apoptosis (Fig.S 2C,C',D). Because the proliferative input of Hh/Dpp signaling is still reduced in these conditions, Ci^{rep} wing size restoration was not complete.

Enrichment of gene functions required for tissue growth

As part of an initial assessment of our screen results we determined the classes of our positive gene set by using *GeneOntology* (GO) terms (Ashburner et al., 2000). GO terms provide a controlled vocabulary to describe gene and gene product attributes for any organism. Each ontology represents a key concept in molecular biology: the molecular function of gene products, their role in multi-step biological processes and their localization to cellular compartments. A gene product can therefore be annotated to several GO terms. For our analysis we defined a gene as *unknown* if it was not associated to any GO term or only to the ontology *Cellular Component*, or if its GO term(s) were only inferred from electronic annotation (GO

evidence code *EA*), and not based on “wet lab” experiments. According to this, our screen uncovered 1’349 unknown genes.

We asked by double-sided statistical tests whether distinct GO terms, i.e. genes with related functions, were significantly over- or underrepresented in our screen hits (Tab. S_). In the phenotype classes affecting wing size (L,S) we found significant enrichment of several terms expected to be linked to tissue growth, such as cell proliferation, cell cycle, growth and cell death. Furthermore we found an enrichment in cell-cell-signaling, cytoskeleton organization and anatomical structure morphogenesis, consistent with a mitotically active, multicellular tissue. Central cellular processes linked to transcription were strongly enriched in the growth phenotype class. However, there was a striking underrepresentation of genes annotated for ribosome/translation. Not unexpectedly, such genes scored mainly in the lethal phenotype class and were significant enriched here. In sum, the analysis of the GO terms validates our screening approach, since we have an enrichment of genes whose GO annotation suggests a role in organ growth. Bioinformatic analysis of GO annotations showed that less than __ of the genes we identified had previously annotated with biological functions consistent with a function in growth. For the majority of our hit genes our data provide a novel functional link to growth regulation.

In the morphology class of wing blisters we found a prevalence of the GO term anatomical structure morphogenesis. A phenotypical hallmark of this class were wings inflated with lymph. Importantly, in the blister class _ genes (_%) have human orthologs. A well known example is *if (inflated)*, which encodes an integrin normally required for cell adhesion-dependent apposition of dorsal and ventral wing blades. Integrin signaling has a strong biological relevance since its disruption is a prerequisite for tumor metastasis (reviewed in Yilmaz and Christofori, 2009). *if* mutant wings have a defective venation causing dehydration and blistering. if^{RNAi} caused strong blistering of the wings in both tester systems, phenocopying the mutant alleles. In addition, we also observed wing size reduction along the D/V axis. In our RNAi screen we found further yet unknown genes to be required for D/V apposition, and are detailed in the Suppl. Info.

Regulators of *Drosophila* wing morphogenesis are enriched for human orthologs

The genes identified in our screen are significantly enriched in human orthologs (1’949 genes; 49%) - 13% more than expected -, potentially indicating a conserved role of these genes for tissue growth also in humans. Among these 50 genes are associated with diseases including cancer (Tab. S3). 6% (225 genes) have no human ortholog, but are conserved in mouse, worm or yeast. 45% (1’805 genes) were fly specific (*Drosophila* and *Anopheles)* with an enrichment of unknown genes (902, including 747 *Drosophila*-only genes) and of genes annotated in chitin metabolism. Among the 1’349 unknown genes were 378 genes with human orthologs with an enrichment for GO terms associated with transcriptional regulation. This group includes *CG6854* and *CG16975* which are detailled below.

CG6854 is a novel regulator of tissue growth

We found *CG6854* as a novel positive regulator of wing growth. Based on two VDRC RNAi lines *CG6854*^{RNAi} strongly reduced the size of the large wing tester below that of wild-type wings (Fig. 5B). The *ci*^{act} wing became roundish-shaped as a result of wing size reduction predominantly along the proximo-distal axis. A third RNAi line (NIG-Fly) caused a similar phenotype (not shown). *CG6854*^{RNAi} expression in otherwise wild-type animals led to wings that were reduced along their PD axes, however, with normal development of the hinge region (Fig. 5C). Additionally, development of marginal wing bristles was disturbed (Fig. 5B,C): long triple-row bristles, specific for the A compartment, appear also on the posterior margin, and the most distal point lacks wing hairs, reminiscent of mild notches. Expression of *CG6854*^{RNAi} in the eye primordium also caused a severe size reduction (Fig. 5D). *CG6854*^{RNAi} expressing clones were indistinguishable from wild-type clones in late 3rd instar wing imaginal discs, but nevertheless gave rise to adult wing phenotypes initially described (Fig. S_). To find out the spatio-temporal requirement of CG6854 we activated *CG6854*^{RNAi} during different larval stages in cell clones and analyzed them in adult wings. To visualize adult clones we co-expressed *forked*^{RNAi} (Fig. S_). We found that CG6854 is predominantly required in the wing marginal region and more distal wing pouch regions (Fig. S_), but not in the proximal wing pouch or wing hinge region (Fig. S_).

We also observed wing size reduction in a background with hyperactivated Hh/Dpp signaling generated by constitutive activity of Smoothened, which causes wing size expansion and ectopic anterior veins (Fig. 5F): *CG6854*^{RNAi} reduced wing extension along the PD axis, but did not alter the Hh-dependent vein defects. In a *nub*::*lgs17E*- or nub::*fringe*- background, where reduced Wg- or Notch-signaling, respectively, cause PD axis reduction based on tissue loss from the wing margins ("notches"), *CG6854*^{RNAi} enhanced the wing margin defect (Fig. 5G,H, and not shown). A cell culture-based RNAi study had suggested *CG6854* could play a positive role in Dpp signaling (Xu et al., 2007). We were curious if this function could account for the phenotypes we had observed. In disagreement with a positive role in Dpp signaling, *CG6854* knockdown did not up-regulate *brk-lacZ* expression (Fig. 5I), which is normally a hallmark of decreased Dpp signaling (Jazwinska et al., 1999; Minami et al., 1999). Moreover, reporters of Hh/Dpp signaling, like *hh-lacZ*, *ptc-lacZ*, *dpp-lacZ*, pMAD or Sal, were not affected (not shown). The wing margin phenotype suggests an interference with signaling pathways required at the distal wing pouch region. However, protein expression at the D/V border required for proximo-distal signaling and growth, like that of Wg and its targets Sens and Vg, or the Notch-signaling target Cut were not affected in wing primordia expressing *CG6854*^{RNAi} (not shown). Solely the Notch-signaling reporter *m8-lacZ* was infrequently and slightly reduced (not shown). Wing size reduction upon *CG6854* depletion is not due to elevated apoptosis, since the level of activated terminal Caspase-3 expression was comparable to control discs (Fig. 5J). In agreement with this finding, the decreased growth phenotype in the wing and the eye could not be ameliorated by co-expression of apoptosis antagonists (not shown). Normal phospho-Histone H3 levels in the wing primordium suggest that cell division (M-phase) was not affected upon *CG6854*^{RNAi} (Fig. 5K). Transition from G1- to S-phase is another important checkpoint during the cell cycle: we found that EdU incorporation during DNA replication (S-phase) was significantly reduced in wing imaginal discs expressing *CG6854*^{RNAi} (Fig. 5L). Consistent

with this, hh>*CG6854*^{RNAi} expression in the P-compartment of imaginal wing discs did cause a compartmental difference of trichome distances in adult wings (not shown). Co-overexpression of cell cycle components CycD/Cdk4 which curb the G1/S transition could ameliorate the nub>*CG6854*^{RNAi} wing phenotype (Fig. 5S), whereas overexpression of String (G2-M transition) had nearly no effect (not shown). Activation of Hippo signaling which normally causes tissue overgrowth by repression of apoptosis and enhanced mitoses was not able to enlarge the size of *CG6854*^{RNAi}-expressing wings (Fig. 5O-R). Curbing mitotic cycles could therefore not rescue or ameliorate the wing size deficit observed in wings whick lack CG6854 function. Wing size reduction upon *CG6854*^{RNAi} is therefore based on impaired cell size, and is independent of mitosis and apoptosis.

CG6854 is conserved from yeast to humans, and encodes three isoforms implicated in transcriptional regulation (Fig. S7): one protein A-isoform with MADF/BESS DNA binding domains, and two highly similar B- and C-isoforms which encode a domain similar to a CTPase. All *CG6854* RNAi lines targeted the B/C-isoforms. Overexpression of the CDS encoding the longest C-isoform did not cause any phenotype in the wild-type wing, but could rescue the wing RNAi phenotype completely (data not shown). We showed that the putative CTPase form of *CG6854* is required for cellular growth, and is specifically required for proximo-distal wing growth. It was previously suggested that Notch signaling cannot solely be integrated by the activity of its nuclear transducer Su(H), but must involve other transcriptional regulators (Cooper, 2000). Based on the Notch-reporter analysis CG6854 might act as a co-regulator of some aspects of Notch signaling.

CG16975/sfmbt is a transcriptional repressor of Wg signaling

We found *sfmbt* in our wing size RNAi screen based on the phenotype “convex-roundish-dusky” which resembled a Wg-signaling gain-of-function. *Sfmbt* remained undetected in cell-based RNAi screens focusing on growth or Wg signaling. The effect of two RNAi lines from the VDRC could be recapitulated by our three costum-made, oligo-nucleotide-based RNAi lines. Overexpression of sfmbt-CDS rescued the sfmbt RNAi phentoype. RNAi of *sfmbt* shows genetic interaction with Wg signaling and caused ectopic activation of Wg-target gene expression (Fig. 6). Because Ubx is ectopically activated by *sfmbt* RNAi in the wing primordium (Fig. 6; *sfmbt mutant:* Hox gene silencing by Sfmbt: Klymenko et al., 2006), we expressed *ubx* RNAi (TRIP and VDRC lines) to see whether the observed phenotype is due to ectopic Ubx. However, knockdown of *ubx* only partially restored the *sfmbt* RNAi wing phenotypes (not shown), indicating that ectopic Wg target expression can be disentangled from ectopic Ubx function. Sfmbt is the *Drosophila* ortholog of a Polycomb-protein-component of a chromatin-regulating complex which conveys transcriptional repression to SUMOylated proteins (Stielow et al., 2008). SUMOylation was previously found to regulate LEF1/TCF, a nuclear effector of mammalian Wnt signaling (Sachdev et al., 2001; Yamamoto et al., 2003). Wnt/Wg signaling is also regulated by another chromatin-regulating complex: the HDAC activity of the NuRD/Hyrax/Paf1 complex, which mediates transcriptional repression, different from the SUMO complex. *hyrax* was previously shown to act as a positive regulator of Wg- and Hh-signaling in

the wing primordium (Mosimann et al., 2006, 2007). dSfmbt and Scm interact synergistically to maintain Polycomb target gene repression (Grimm et al., 2009: Abd-B and En are ectopically expressed only if both Sfmbt and Scm are lost; Oktaba et al., 2008: en is misexpressed in many dSfmbt mutant clones; Dll expression is lost in Scm single mutant clone (our *sfmb* RNAi experiments: Abd-A-lacZ and Abd-B-lacZ: no ectopic expression).

Comparison of the *in vivo* wing growth screen to cell culture based RNAi screens

Exclusiveness & overlap (GO enrichments and depletions): wing screen vs. cell culture growth screens (9 *Drosophila* +/- 1 *Human*)

In recent years many large scale RNAi studies have been undertaken *in vitro* – in cultured *Drosophila* cell lines - to analyse cell cycle, cell growth and cell survival (reviewed in Cully and Leevers, 2006; Pollard, 2003). Since our screen addressed some of the same processes but in an *in vivo* context we were interested in systematically comparing our results (Fig. S5, blue circle) with a selection of nine RNAi screens focused on cell-cycle/-growth associated phenotypes, which were based on FACS- and imaging-supported cell culture RNAi assays (Venn-Diagram Fig. S5, red circle; Björklund et al., 2006; Boutros et al., 2004; Bettencourt-Dias et al., 2004, Kiger et al., 2003; Eggert et al., 2004; Echard et al., 2007; Goshima et al., 2007; Yu et al., 2007; Lu et al., 2007). Clearly, the comparison is somewhat arbitrary since it is based on screens with different read-outs and with different "cut-offs" for genes defined as hits, influencing the degree of overlap. In any case, from our total positive gene set (3'979, including off targets) we found 471 genes in common with the growth-based culture screens (939 genes). The overlap included mostly known genes, and only 6.7% (32) unknown genes. In contrast, the fraction uniquely found in our screen is highly enriched for unknown genes (37.5% or 1'317 genes). Notably, most genes of this unique wing fraction were conserved during evolution up to higher vertebrates. We were interested to see if we could further distinguish the 'uniquely-wing' from 'uniquely-culture' based gene sets.

Genes only found in our screen are enriched for GO terms linked to lipid metabolism, whereas culture based screens show no such enrichment (Tab. S4). One example to explain this finding might be Hh signaling itself, because the activity of Hh depends on a direct dual lipidation, and our tester systems are sensitive to factors which modulate the strength of the Hh signal. Since the Hh protein is lipidated by a palmitoyl and cholesterol moiety, depleting factors involved in lipid/cholesterol absorption likely affect the ci^{act} tester. An example of a gene we identified in this class was *NPC1b*. $NPC1b^{RNAi}$ significantly reduced the large wing but not in the small wing tester. So far, no mutant phenotype is available for *NPC1*. The NPC1 family of proteins plays crucial roles in the intestinal absorption and intracellular trafficking of sterols. Another significantly enriched functional class concerns carbohydrate metabolism. The functions of many extracellular proteins, including those of HSPGs, rely on sugar modifications. ECM molecules often have auxiliary functions in signaling and are likely functionally redundant, therefore knockdown of such components does not always lead to a strong phenotype, if at all. This is particularly the case in cell

culture RNAi screens based on autocrine signaling, since ECM components are required for cell-to-cell signaling, e.g. for the establishment of morphogen gradients. Likewise, a RNAi phenotype of well known carbohydrate ECM components such as *boi, botv, iHog, dally or dlp* which scored positive in our *in vivo* screen (Tab. S1, Tab. S2), remained largely undetected in most in cell culture screen. Components of the chitin metabolism also belong to the carboyhydrate metabolism class. Two examples of genes involved in chitin metabolism are $CG5883^{RNAi}$ and $CG7017^{RNAi}$ caused wrinkled wing epithelia as well as reduced wing sizes. No mutant phenotype has so far been reported for these genes, however, *CG5883* RNAi led to impaired secretory activity of S2 cells (Bard et al., 2007). A chitin-based extracellular matrix not only stabilizes the exoskeleton but might also play a role in growth (Kawamura at el., 1996). Likewise, the secreted ZP (Zona pellucida) proteins Miniature and Dusky link the ECM with the cytoskeleton and are required for proper wing growth (Roch et al., 2003). A third, highly enriched GO term of genes only found in our screen is transmembrane transporter/receptor activity. Although not statistically enriched, our screen detected more frequently genes encoding components of the cytoskeleton. For example, RNAi against genes encoding intracellular adaptors of the cytoskeleton like *wasp*, *zyxin* or *chi* reduced the tester wing sizes, but did not score in cell culture based screens.

Of emerging biological interest are kinases and phosphatases**.** They are complementary regulators of protein activity and comprise a relatively large protein class within the proteome. Phosophorylation is a fundamental strategy used by eukaryotic organisms to regulate a host of biological functions, including DNA replication, cell cycle progression, energy metabolism, and cell growth and differentiation. Many human tumors carry genetic alterations in at least one phosphatase or kinase gene (Futreal et al., 2005; Greenman et al., 2007). 236 kinases and 84 phosphatases are predicted in the *Drosophila* protein set (kinase.com; Morrison et al., 2000). 200 kinases and 78 phosphatases were represented in the RNAi library. Our screen identified 76 kinases and 26 phosphatases as regulating growth. Kinases are central to Hh signaling. For example, Pka-C1 negatively or positively influences Hh signaling, depending on the dose (Zhou et al., 2006). RNAi against Pka-C1 reduced size in both wing conditions. Our screen results share 56 kinases and 16 phosphatases with different culture based RNAi screens mentioned above, including one focusing only on kinases (Bettencourt-Dias et al., 2004), supporting the notion that kinases can also be involved in the crosstalk of different signaling pathways (Cully et al., 2006; Katz et al., 2007). 21 kinases and 19 phosphatases were found only by our wing screen.

We scored the highly conserved gene *CG3534*, which encodes a putative carbohydrate kinase, as a growth repressor: $CG3534^{RNAi}$ caused variable enlargement of small ci^{rep} wings, although predominantly along their PD axes, and blistering of large ci^{act} wings. No mutant allele or RNAi phenotype had previously been reported for this kinase. A further promising example is the tumor suppressor ASPP, which regulates C-terminal Src kinase activity (Langton et al., 2007): $ASPP^{RNAi}$ caused enlarged wings in both of our tester conditions. It has not been described in cell culture RNAi screens. A similar phenotype (enlargement in ci^{act} wings) was also seen for *CG3216,* an unknown gene which encodes a putative GPCR-coupled

kinase. In cultured fly cells its depletion led to aberrant mitosis (Bettencourt-Dias et al., 2004); no phenotype based on a mutant allele is currently reported. Consistent with our observations that it is a putative growth repressor *CG3216* is homologous to the human Atrial natriuretic peptide receptor, which inhibits cell proliferation (Kong et al., 2008). As a positive growth regulator we found *lal/AuroraB* kinase, whose knock-down caused wing size reduction in the ci^{act} condition (pupal death in ci^{rep} wings). No mutant allele is available so far; however, RNAi in cultured cells caused cell number decrease in the G1 phase and a cell number increase in the G2 phase of the cell cycle (Björklund et al., 2006), supporting the aberrant cytokinesis phenotypes observed in previous RNAi studies (Lange et al., 2002; Giet and Glover, 2001). We found *mts (microtubule star)*, which encodes the PP2A subunit of a phosphatase complex, as a suppressor of the large ci^{act} tester wing. PP2A is required for cell polarity/adhesion, ommatidial differentiation, mitotic spindle organisation and cell proliferation (Dombradi et al., 1990; Kiger et al., 2003; Eggert et al., 2004). RNAi in *Drosophila* tissue culture revealed a positive function in Hh signaling (Nybakken et al., 2005; G.H. and K.B., unpubl.), supported *in vivo* by a *smo* mutant modifier screen (Casso et al., 2008).

Pooling all results of the cell-based RNAi screens and comparing this with our screen generates a set of __ genes that potentially regulate cell growth / imaginal disc growth (Fig. S5). This leaves __ genes uniquely identified in our screen. The overlap, however, among the cell-culture screens (__ genes) or the total overlap together with our in vivo screen (__ genes) is not very extensive, and unknown genes are underrepresented: we addressed only 83% of the genome as compared to culture based screens, which addressed altogether the entire genome. The occurrence of false negative results might be higher in our wing screen based on the nature of the RNAi source: i) P-element transgenes are affected by genomic position effects, ii) a gene was often addressed only by one RNAi line, and iii) on average 300 nucleotides of the coding region was used for VDRC siRNA design, whereas RNAi of cell culture screens often targeted larger parts and at least 2 independent regions of the coding sequence (www.flyrnai.org, www.genomernai.org). In culture based screens, cellular growth on one hand, and signaling cascades on the other hand have been addressed by functional readouts primarily based on the analysis of individual cells. The mostly cell autonomous read-outs did not allow the detection of components normally required for a multicellular tissue to interact at the cellular level and to grow. Genes may also encode proteins whose function are required at the interface of two or more pathways or whose function are only needed for organ growth. In contrast**,** our *in vivo* screen targets tissue growth and patterning at the same time.

Why an in *in vivo* screen is superior: cell-to-cell signaling is central to intact organ growth

One important advantage of our *in vivo* growth screen over cell-based screens is the detection of tumor suppressor gene functions. An example how a TSG's function is coupled to tissue growth regulators is the ubiquitin ligase Hyperplastic discs (Hyd), which negatively regulates *hh* and *dpp* expression by independent mechanisms and also impacts on the stability of Ci (Lee et al., 2002). We scored hyd^{RNAi} as giving severely blistered wings in the ci^{act} condition (Fig. S1 J). Furthermore, components of the Insulin- or

Hippo (Hpo) signaling pathways, which were readily detected in both of our wing conditions, remained largely undetected in cell culture-based RNAi screens. Discovered initially in *Drosophila*, the Hpo pathway has emerged as an important conserved signaling pathway that controls organ size during development by restricting cell growth and proliferation and by promoting apoptosis (Harvey et al., 2003). Abnormal activity the Hpo pathway has been implicated in human cancer (Harvey and Tapon, 2007). Of the 10 established components of this pathway we found 7 (*ex*, *mer*, *ds*, *d, ft, sav* and *wts* (Fig. 2, Fig. S1, Tab. S1). RNAi against growth restricting components of the pathway varied in effects: depletion of *mer*, *warts* and *expanded* caused enlargement of wings in both testers; in contrast, targeting the atypical cadherins encoded by *fat* and *dachsous* caused enlargement only in the small wing tester. In the large-wing tester depletion of these gene functions caused predominantly epithelial blistering and more folds, or necrosis, respectively. Of note, our screen set-up did not allow us to test for overgrowth effects observed in systemic TSG mutants that are based on a developmental extension of the larval period (Bilder, 2004), since *nub*-Gal4 is not active in the entire animal.

An important control function for the integration of different pathway activities is exerted by the ESCRT machinery (endosomal sorting complex required for transport), which mediates sorting of ubiquitinated membrane proteins, like ligand-activated receptor tyrosine kinases, into the intraluminal vesicles of multivesicular endosomes, thereby targeting them for lysosomal degradation in lysosomes. Clonal loss of *dTsg101/erupted/vps23* or *vps25*, which encode components of the ESCRT I and II complex, lead to overgrowth involving the activation of the Notch- and the Jak/Stat pathways, caused by impaired downregulation of the Notch receptor (Moberg et al., 2005; Thompson et al., 2005).

Wing growth is repressed by *CG14542/vps2*

We found the highly conserved gene *CG14542/vps2* to act as a negative and novel regulator of wing growth. $CG14542^{RNAi}$ caused significant enlargement in the small wing condition (Fig. 7B,C), and pupal lethality in the ci^{act} as well as in the wild type background, and it also caused enlarged eyes (not shown; Fig. 7J, Fig. S6). Milder ubiquitous $CG14542^{RNAi}$ expression in the wing primordium resulted in overgrown wings with thickened veins (Fig. 7D,E; Fig. S6). As a similar effect is also observed upon increased *CG14542* expression (Fig. 7R → *into Supplement*), it is likely that *CG14542* acts in a dominant negative manner when over-expressed. Normal spacing of adult wing blade trichomes and elevated *string* mRNA levels in $CG14542^{RNAi}$ wing primordia (Fig. S6; not shown) suggest that CG14542 controls cell number, but not cell size. We also observed sensory organ defects indicative of reduced Wg/Notch signaling or elevated Hh signaling (Fig. S6). To test the hyperproliferative behavior we made use of $Minute^{+/-}$ clones, which would normally under-grow and die (Morata et al., 1975). $Minute^{+/-}$ clones were rescued and even over-grew by activating pro-growth pathways like Hh/Dpp signaling within such clones (not shown). Unexpectedly, $Minute^{+/-}$ clones could not be rescued by $CG14542^{RNAi}$ within such clones, and were even more rapidly lost from the wing primordium (not shown). When clones of $CG14542^{RNAi}$ were generated in the wild-type background, we observed roundish-shaped clones with pyknotic, dying cells indicating cell

death. Terminal caspase-3 activation, a prerequisite for apoptosis, was highly activated in *CG14542*^{RNAi} clones (Fig. 7K,K', close-up of a pouch clone). Clones died when they are located within the wing pouch, but they round up and overgrow outside of the pouch (Fig. 7L). We found several pro-growth signaling pathways activated in clones inside (dying) and clones outside (overgrowing) of the wing pouch: pMad, the nuclear transducer of Dpp signaling and its target gene *sal* (Fig. 7N, Fig. S6A, and not shown), Wg (Fig. 7O), and Upd, a Notch target gene which encodes the ligand of Jak/Stat signaling (not shown). Consistent with the upregulation of these pro-growth cues, global over-expression of *CG14542*^{RNAi} led to enhanced phospho-Histone expression (Fig. 7P). However, signaling pathways upregulated in the clonal condition were mostly unaltered in wings globally expressing *CG14542*^{RNAi}, and pMad domains were distorted, likely due to the strong folding of the overgrown wing disc (Fig. S6B; Ptc, Upd: not shown). Among all tested markers we only found Wg epression altered. The ring of Wg expression that coincides with the most proximal wing and the most distal hinge region was extended (Fig. 7Q-Q''), and *nkd*, a Wg target, was upregulated as assessed by quantitative RT-PCR (not shown). Strong and expanded Wg expression was already apparent during early L3 stages, where normally ring expression is only faintly occurring in a slim ring. Of note, 30A-Gal4 driven expression of *CG14542*^{RNAi} in the ring domain could recapitulate the overgrowth effect (not shown). In contrast, Wg expression at the D/V border was not altered, and Vg expression, which requires Wg D/V expression, was also normal (not shown). Initiation of Wg ring expression normally requires the confrontation of Vg-positive (pouch) versus Vg negative (hinge) tissue, whereas *el* and *noc* repress Wg ring expression (Weihe et al., 2004). *(wg transcription / wg-lacZ affected?).*

We associate the overgrowth effect of global *CG14542* knockdown, which is predominantly enhanced along the distal-proximal direction (Fig. 7P,Q,Q'), with elevated Wg ring expression at the interface of the wing pouch/hinge region. Ablation of Wg ring expression in *spade* mutants causes reduced wing size predominantly along the PD direction (Tiong and Nash, 1990; Couso et al., 1994). The cause for deregulated proximo-distal growth in the global knockdown likely differs from local overgrowth and upregulation of several growth factors in the clonal scenario. In the latter case, regions outside the pouch are competent for overgrowth, whereas potentially overgrowing clones within the pouch are counteracted by simultaneous autonomous activation of apoptosis. Understanding the cause for the selective competence of potentially overgrowing tissue requires further investigation. *CG14542* plays a putative role in protein transport in the ESCRT III subcomplex of the ESCRT/Vps machinery based on electronic annotation. Previously, a *CG14542* MENE mutant (mutant eye disc no eclosion) was reported to cause pupal death (Menut et al., 2007) and intracellular Notch accumulation in eye disc clones, reminiscent of other ESCRT components (Vaccari et al., 2009). However, the previous studies were limited to clonal assays due to systemic lethality. In contrast, we could show an additional and, likely, alternative way for CG14542/Vps2 to elicit overgrowth upon RNAi in a non-clonal condition. Further, the protein interaction database STRING reveals an interaction with *Chmp1/CG4108*, which we also scored as a negative growth regulator in our screen (Fig. S7). Similar to the depletion of *CG14542*, *Chmp1*^{RNAi} enhanced the *ci*^{rep} wing

size. No mutant phenotype is known of *Chmp1*, and neither *CG14542* nor *Chmp1* scored positive in any cell culture RNAi screen focusing on cell cycle or signaling pathways.

Conclusion

Our search for growth regulators exemplifies the application of functional genomics in developmental biology. We interrogated the genome of *D. melanogaster* by transgenic RNAi by performing an *in vivo* RNAi screen which allowed us to generate an extensive catalogue of genes required for tissue specific growth regulation. The inducible system used in our RNAi screen further allowed to detect gene functions which would have remained uncovered by classical mutagenesis due to early developmental lethality. The screen recovered most of the predicted genes required for growth and related processes. Importantly, our study identified novel regulators of growth and morphology of the *Drosophila* wing epithelium. By monitoring growth within an entire organ analysis of the candidates will help us to understand in future how growth control of a complex tissue is achieved. We identified 2'596 genes (incl. off-target prone genes) of as being involved in growth control, including 1'349 genes with previously unascribed function. Furthermore, genes with human orthologs are significantly enriched. The analysis of *CG6854/CTPsyn*, *CG16975*/*sfmbt* and *CG14542/vps2* indicate how our tissue specific *in vivo* RNAi study complements cell culture-based screens which looked at individual cells and reporter-based detection of signaling activities uncoupled from entire organs, thereby often fail to detect certain gene function requirements. In ongoing and future studies on candidates from our screen we aim to uncover how morphogenetic pathways connect ultimately to components of the cell cycle machinery. An important additional step will be to determine how orthologs of the components identified here link to molecular pathways of higher vertebrates, and thereby further our understanding of diseases caused by deregulated growth.

Acknowledgements

Krystyna Keleman and the members of the VDRC library for supporting G.R. with RNAi lines during the screen

Barry Dickson and his laboratory for hosting G.R. during the RNAi screen

Jennifer Mummery-Widmer for help with FileMaker (Jürgen Knoblich lab)

Esther Jud for help with fly stock propagation

Manuel Stark for assistance with the STRING database (Christian v. Mering lab)

George Hausmann for critical input on the manuscript

References

Aegerter-Wilmsen, T., et al. (2007). Model for the regulation of size in the wing imaginal disc of *Drosophila*. Mech. Dev. 124, 318-326

Affolter, M., and Basler, K. (2007). The Decapentaplegic morphogen gradient: from pattern formation to growth regulation. Nat Rev Genet. 8(9), 663-674

Ashburner, M., et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 25, 25-29

Baeg, G.H., et al. (2007). Genome-wide RNAi analysis of JAK/STAT signaling components in *Drosophila*. Genes & Dev. 19, 1861-1870

Bard, F., et al. (2006). Functional genomics reveals genes involved in protein secretion and Golgi organization. Nature 439, 604-607

Bentley, A.M., et al. (2002). Phenotypic characterization of *Drocophila ida* mutants: defining the role of APC5 in cell cycle progression. J.Cell.Sci 115, 949-961

Bettencourt-Dias, M., et al. (2004). Genome-wide survey of protein kinases required for cell cycle progression. Nature 432, 980-987

Bilder, D. (2004). Epithelial polarity and proliferation control: links from the *Drosophila* neoplastic tumor suppressors. Genes Dev 18, 1909-1925

Billin, A.N., et al. (1991). Isolation of a family of *Drosophila* POU domain genes expressed in early development. Mech Dev. 34(2-3), 75-84

Bischof, J., et al. (2007). integrasesAn optimized transgenesis system for Drosophila using germ-line-specific ΦC31. PNAS 104(9), 3312-3317

Björklund, M., et al. (2006). Identification of pathways regulating cell size and cell-cycle progression by RNAi. Nature 439, 1009-1013

Boutros, M., et al (2004). Genome-Wide RNAi Analysis of Growth and Viability in *Drosophila* Cells. Science 203, 832-835

Brand, A. H. and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. Development 118, 401-415

Britton, J.S, and Edgar, B.A. (1998). Environmental control of the cell cycle in Drosophila: nutrition activates mitotic and endoreplicative cells by distinct mechanisms. Development 125(11), 2149-2158

Bryant, P.J., and Simpson, P. (1984). Intrinsic and extrinsic control of growth in developing organs. Q Rev Biol. 59(4), 387-415

Calleja, M., et al. (1996). Visualization of gene expression in living adult *Drosophila*. Science 274 (5285), 252-5

Campbell, G., and Tomlinson, A. (1999). Transducing the Dpp morphogen gradient in the wing of Drosophila: regulation of Dpp targets by brinker. Cell 96, 553-562

Casso, D.J., et al. (2008). A screen for modifiers of hedgehog signaling in *Drosophila* melanogaster identifies swm and mts. Genetics 178, 1399-1413

Cifuentes, F.J., and Garcia-Bellido, A. (1997). Proximo-distal specification in the wing disc of *Drosophila* by the *nubbin* gene. Proc Natl Acad Sci U S A. 94(21), 11405-11410

Citnan et al. (2007). Protothoracic gland

Cook, O., et al. (2004). *brinker* and *optomotor-blind* act coordinately to initiate development of the L5 wing vein primordium in *Drosophila*. Development 131, 2113-2124

Cook, M., and Tyers, M. (2007). Size control goes global. Curr Opin Biotechnol. 18(4), 341-350

Couso, J.P., et al. (1994). The Wingless signalling pathway and the patterning of the wing margin in *Drosophila*. Development 120(3), 621-636

Clark, L., et al. (2004). Prothoracic gland

Cooper, M.T., et al. (2000). Spatially restricted factors cooperate with Notch in the regulation of Enhancer of split genes. Dev Biol. 221(2), 390-403

Crickmore, M.A., and Mann, R.S. (2006). Hox control of organ size by regulation of morphogen production and mobility. Science 313, 63-68

Crozatier, M., et al. (2004). Patterns in evolution: veins of the Drosophila wing. Trends Genet. 20(10), 498-505

Cully, M.J., and Leevers, S.J. (2006). RNA interference pinpoints regulators of cell size and the cell cycle. Genome Biol. 7(5), 219

DasGupta, R., et al., (2005). Functional Genomic Analysis of the Wnt-Wingless Signaling Pathway. Science 308, 826-833

DasGupta, R., and Gonsalves, F.C. (2008). High-throughput RNAi screen in *Drosophila*. Methods Mol Biol. 469, 163-84

de Celis, J.F., et al. (1996). Activation and function of Notch at the dorsal-ventral boundary of the wing imaginal disc. Development 122, 359-359

Diaz-Benjumea, and F.J., Cohen, S.M. (1993). Interaction between dorsal and ventral cells in the imaginal disc directs wing development in *Drosophila.* Cell 75, 741-752

Dick, T., et al. (1991). Two closely linked *Drosophila* POU domain genes are expressed in neuroblasts and sensory elements. Proc Natl Acad Sci U S A. 88(17), 7645-7649

Dietzl, G., et al. (2007). A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. Nature 448, 151-156

Dombradi, V., et al., (1990). Protein phosphatase 1 activity in *Drosophila* mutants with abnormalities in mitosis and chromosome condensation. FEBS Lett. 275, 39-43

Echard, A., et al. (2004). Terminal Cytokinesis Events Uncovered an RNAi Screen. Curr. Biol. 14, 1685-1693

Echeverri, C.J, and Perrimon, N. (2006). High-throughput RNAi screening in cultured cells: a user's guide. Nat Rev Genet. 7(5), 373-84

Eggert, U.S., et al. (2004). Parallel Chemical Genetic and Genome-Wide RNAi Screens Identify Cytokinesis Inhibitors and Targets. PLOS 2, 2135-2143

Felsenfeld, A.L. and Kennison, J.A. (1995). Positional signaling by *hedgehog* in *Drosophila* imaginal disc development. Development 121, 1-10

Fuchs, F., and Boutros, M. (2006). Cellular phenotyping by RNAi. Brief Funct Genomic Proteomic. 5(1), 52-6.

Futreal, P.A., et al. (2005). Somatic mutations in human cancer: insights from resequencing the protein kinase gene family. Cold Spring Harb Symp Quant Biol. 70, 43-49

Garcia-Bellido, A., and Dapena, J. (1974). Induction, detection and characterization of cell differentiation mutants in Drosophila. Mol Gen Genet. 128(2), 117-30

Giet, R., and Glover, D.M. (2001). *Drosophila* aurora B kinase is required for histone H3 phosphorylation and condensin recruitment during chromosome condensation and to organize the central spindle during cytokinesis. J Cell Biol. 152(4), 669-682

Goshima, G., et al. (2007). Genes Required for Mitotic Spindle Assembly in *Drosophila* S2 cells. Science 316, 417-421

Greenman, C., et al. (2007). Patterns of somatic mutation in human cancer genomes. Nature 446 (7132), 153-158

Guo, Y., et al. (2008). Functional genomic screen reveals genes involved in lipid-droplet formation and utilization. Nature 453 (7195), 657-61

Haley, B., et al. (2008). A simplified miRNA-based gene silencing method for *Drosophila melanogaster*. Dev Biol. 321(2), 482-490

Huangfu, D., and Anderson, K.V. (2006). Signaling from Smo to Ci/Gli: conservation and divergence of Hedgehog pathways from *Drosophila* to vertebrates. Development 133, 3-14

Hufnagel, L., et al. (2007). On the mechanism of wing size determination in fly development. Proc. Natl. Acad. Sci. USA 104, 3835-3840.

Jazwinska, A., et al. (1999). The Drosophila gene *brinker* reveals a novel mechanism of Dpp target gene regulation. Cell 96, 563-573

Katz, M., et al. (2007). Regulation of MAPKs by growth factors and receptor tyrosine kinases. Biochim Biophys Acta. 1773(8), 1161-76

Kong, X., et al., (2008). Natriuretic peptide receptor a as a novel anticancer target. Cancer Res. 68,249-56

Lee, L.A., et al. (2001). A genetic screen for suppressors and enhancers of the *Drosophila* pan gu cell cycle kinase identifies cyclin B as a target. Genetics 158, 1545-1556.

Lents, N.H., and Baldassare, J.J. (2006). RNA interference takes flight: a new RNAi screen reveals cell cycle regulators in *Drosophila* cells. Trends Endocrinol Metab. 17(5), 173-4

Leopold, P., and Perrimon, N. (2007). Drosophila and the genetics of the internal milieu. Nature 450(7167), 186-188

Lloyd, A., and Sakonju, S. (1991). Characterization of two *Drosophila* POU domain genes, related to *oct-1* and *oct-2*, and the regulation of their expression patterns.

Janody, F., et al. (2003). Two subunits of the *Drosophila* mediator complex act together to control cell affinity. Development 130, 3691-3701

Kawamura et al. (1996).

Kawamura et al. (1999).

Kiger, A.A., et al. (2003). A functional genomic analysis of cell morphology using RNA interference. J. Biol. 2(4):27

Lange, B.M., et al. (2002). Cdc37 is essential for chromosome segregation and cytokinesis in higher eukaryotes. EMBO J. 21(20), 5364-5374

Langton, P.F., et al. (2007). Drosophila ASPP regulates C-terminal Src kinase activity. Dev Cell. 13(6), 773-82

Lum, L., et al. (2003). Identification of Hedgehog Pathway Components by RNAi in *Drosophila* Cultured Cells. Science 299, 2039-2045

Martin, F.A., et al. (2004). *brinker* gradient controls wing growth in *Drosophila*. Development131, 4921-4930

Mathey-Prevot, B., and Perrimon, N. (2007). Matter arising: off-targets and genome-scale RNAi screens in *Drosophila*. Fly 1(1), 1-5

Menut, L., et al. (2007). A Mosaic Genetic Screen for *Drosophila* Neoplastic Tumor Suppressor Genes Based on Defective Pupation. Genetics 177, 1667-1677

Methot, N., and Basler, K. (1999). Hedgehog controls limb development by regulating the activities of distinct transcriptional activator and repressor forms of Cubitus interruptus. Cell 96(6), 819-831

Methot, N., and Basler, K. (2000). Suppressor of Fused opposes Hedgehog signal transduction by impeding nuclear accumulation of the activator form of Cubitus interruptus. Development 127, 4001-4010

Minami, M., et al. (1999). *brinker* is a target of Dpp in Drosophila that negatively regulates Dpp-dependent genes. Nature 398, 242-246

Marty, T., et al. (2000). Schnurri mediates Dpp-dependent repression of brinker transcription. Nat. Cell Biol. 2, 745-749

Morata, G., and Ripoll, P. (1975). Minutes: mutants of drosophila autonomously affecting cell division rate.

Dev Biol. 42(2), 211-221

Morrison, D.K., Murakami, M.S., Cleghon, V. (2000). Protein kinases and phosphatases in the *Drosophila* genome. J Cell Biol. 150(2), F57-62

Müller, D., et al. (2005). Genetic Modifier Screens on Hairless Gain-of-Function Phenotypes Reveal Genes Involved in Cell Differentiation, Cell Growth and Apoptosis in *Drosophila* melanogaster. Genetics 171, 1137-1152

Müller, P., et al. (2005). Identification of pathways regulating cell size and cell-cycle progression by RNAi. Nature 436, 871-875

Natzle, J.E., et al. (2008). Bursicon signaling mutations separate the epithelial-mesenchymal transition from programmed cell death during *Drosophila melanogaster* wing maturation. Genetics 180(2), 885-93

Neto-Sliva, R.M., et al., (2009). Mechanisms of growth and homeostasis in the Drosophila wing. Annu Rev Cell Dev Biol. 25, 197-220

Ng, M., et al. (1996). Nubbin encodes a POU-domain protein required for proximal-distal patterning in the Drosophila wing. Development 121(2), 589-599

Nybakken, K., et al. (2005). A genome-wide RNA interference screen *Drosophila melanogaster* cells for new components of the Hh signaling pathway. Nat. Gen. 37, 1323-1332

Ogden, S.K., et al. (2008). G protein Galpha-i functions immediately downstream of Smoothened in Hedgehog signalling. Nature. 2008 456(7224), 967-70

Peabody, N.C., et al. (2008). Bursicon functions within the *Drosophila* CNS to modulate wing expansion behavior, hormone secretion, and cell death. J Neurosci. 28(53), 14379-14391

Perrimon, N., and Mathey-Prevot, B. (2007). Matter arising: off-targets and genome-scale RNAi screens in *Drosophila*. Fly 1(1), 1-5

Pollard, T.D. (2003). Functional genomics of cell morphology using RNA interference: pick your style, broad or deep. J Biol. 2(4), 25

Reddy, B.V., and Irvine, K.D. (2008). The Fat and Warts signaling pathways: new insights into their regulation, mechanism and conservation. Development 135(17), 2827-38

Reis and Edgar (2004). Negative regulation of dE2F1 by cyclin-dependent kinases controls cell cycle timing. Cell 117(2), 253-264

Rong, Y.S., et al. (2002). Targeted mutagenesis by homologous recombination in D. melanogaster. Genes Dev. 15;16(12), 1568-81

Saucedo, L.J., and Edgar, B.A. (2007). Filling out the Hippo pathway. Nat Rev Mol Cell Biol. 8(8), 613-21

Schwank, G., et al. (2008). Growth regulation by Dpp: an essential role for Brinker and a non-essential role for graded signaling levels. Development 135, 4003-4013

Slusarski, D.C., et al. (1995). Mutations that alter the timing and pattern of cubitus interruptus gene expression in Drosophila melanogaster. Genetics 139(1), 229-240

Somna, M.P., et al. (2002). Molecular Dissection of Cytokinesis by RNA Interference in *Drosophila* Cultured Cells. Mol.Biol.Cell 13, 2448-2460

Steinbrink, S., and Boutros, M. (2008). RNAi screening in cultured *Drosophila* cells. Methods Mol Biol. 420, 139-153

Strutt, H. (2005). Organ shape: controlling oriented cell division. Curr Biol. 15(18), R758-9

Tabata, T., and Takei, Y. (2004). Morphogens, their identification and regulation. Development. 131, 703-12

Tiong, S.Y., and Nash, D. (1990). Genetic analysis of the adenosine3 (Gart) region of the second chromosome of *Drosophila melanogaster*. Genetics 124(4), 889-897

Tönissoo, T., et al. (2006). Heterozygous mice with Ric-8 mutation exhibit impaired spatial memory and decreased anxiety. Behav Brain Res. 167(1), 42-48

Tyler, D.M., et al. (2007). Genes affecting cell competition in Drosophila. Genetics 175(2), 643-57

Vaccari, T., et al. (2009). Comparative analysis of ESCRT-I, ESCRT-II and ESCRT-III function in *Drosophila* by efficient isolation of ESCRT mutants. J Cell Sci 122 (14): 2413-2423

Watts, R.J., et al. (2003). Axon pruning during *Drosophila* metamorphosis: evidence for local degeneration and requirement of the ubiquitin-proteasome system. Neuron 38, 871-885.

Wang, H., et al. (2005). Ric-8 controls *Drosophila* neural progenitor asymmetric division by regulating heterotrimeric G proteins. Nat. Cell. Biol. 7, 1091-98

Weihe, U., et al. (2004). Proximodistal subdivision of *Drosophila* legs and wings: the *elbow-no ocelli* gene complex. Development 131(4), 767-774

Xu, L., et al. (2007). Msk is required for nuclear import of TGF-β/BMP- activated Smads. JCB 178, 881-884.

Yi, C.H., et al. (2007). A genome-wide RNAi screen reveals multiple regulators of caspase activation. JCB 179, 619-626

Yilmaz, M., Christofori, G. (2009). EMT, the cytoskeleton, and cancer cell invasion. Cancer Metastasis Rev. 28(1-2), 15-33

Zhang, S.L., et al (2006). Genome-wide RNAi screen of Ca(2+) influx identifies genes that regulate Ca(2+) release-activated Ca(2+) channel activity. Proc Natl Acad Sci U S A. 103(24), 9357-9362

Figure Legends

Figure 1. The experimental setup of the growth RNAi screen
Two different forms of Ci, the transcriptional mediator of Hh signaling, were used to generate the tester fly strains: *ci^{rep}* acts as a transcriptional repressor of Hh-dependent target genes, *ci^{act}* is a constitutive activator of Hh-dependent target expression like *ptc* or *dpp.* Increase or decrease of Hh signaling in the wings of the two different tester strains leads to stereotyped size alterations and vein abnormalities. The VDRC RNAi library was assayed with a large-wing (*ci^{act}*) and a small-wing (*ci^{rep}*) tester fly strain. *nub-Gal4* activates expression of the RNAi transgenes in the larval wing pouch, which develops into the adult wing (orange oval). The adult F1 progeny was scored for suppression or enhancement of tester wing sizes and for alterations of epithelial wing morphology.

Figure 2. Efficient depletion of gene function required during wing development
(A) Wild-type wing. (B) *ci^{act}* tester wing. (B') *ci^{rep}* tester wing. (C) *smo^{RNAi}* suppress the ci^{act} and (C') enhances the *ci^{rep}* phenotype. (D) *ci^{RNAi}* reduces *ci^{act}* wings, and (D') only slightly affects *ci^{rep}* wings. (E,E') *pten^{RNAi}* enlarges *ci^{act}* and *ci^{rep}* wings without altering the Hh signaling dependent patterning defects. (F,F') *mer^{RNAi}* strongly enlarges *ci^{act}* and *ci^{rep}* wings, accompanied by severe folding of the wing cuticle.

Figure 3. Results overview of the RNAi screen
(A) A total of 11'606 (11'466 genes) *Drosophila* protein coding genes, addressed by 21'014 (21'206) RNAi lines, was analyzed. RNAi of 3.979 genes led to an adult wing phenotype, including 654 off target-prone genes (hatched segment). (B) Main phenotype classes found in the screen are alterations of wing size (S/smaller wings, L/larger wings, R/roundish wings), morphology and lethality (hatched segments and grey numbers: off target-prone genes).

Figure 4. Classes of predominant wing phenotypes scored in the RNAi screen
(A) Wild-type wing. (B) *ci^{act}* tester wing. (B') *ci^{rep}* tester wing. (D-D''') Different strengths in size reduction of *ci^{act}* tester wings upon RNAi against different genes. (E,E') Wing size enlargement upon RNAi against two different genes. (F,F') Round shaped wing phenotypes primarily based on reduction along the proximo-distal axis. (G,G') Wing blistering in *ci^{act}* wings. The dusky wing surface (G) indicates an aberrantly dense trachoma formation. (G'') Impaired wing unfolding. (H) Vein reduction in *ci^{act}*. (H') Wing trichoma defect in *ci^{act}* accompanied by size reduction and broadened veins.

Figure 5. *CG6854* is a novel gene required for proximo-distal wing growth
(A) Wild-type wing. (B) *CG6854^{RNAi}* causes significant size reduction of *ci^{act}* tester wings, predominantly along the PD axis. (C) *CG6854^{RNAi}* causes a less strong size defect in the wild-type background. (D) Eye size reduction upon *CG6854^{RNAi}* (left inset) in comparison to a control eye (right inset). (E,F) PD shortening upon *CG6854^{RNAi}* is apparent in sal::*smo^{1}* wings. (G,H) *CG6854^{RNAi}* does not cause further PD shortening or enhancement of notches in sal::*lgs^{17E}* wings. (I) *brk* reporter gene expression is upregulated upon *CG6854^{RNAi}* (left inset). (J) *m8* reporter expression is lost in the posterior wing discs compartment upon expression of *CG6854^{RNAi}* by hh-Gal4 in the posterior wing compartment. (K) pHis expression is reduced by *CG6854^{RNAi}* expression in the wing pouch (left inset, compare to non-pouch regions with normal pHis expression).

Figure 6. *sfmbt* is a repressor of Wg signaling

Figure 7. *CG14542* is a novel tumor suppressor
(A) Wild-type wing. (B,C) *CG14542^{RNAi}* causes enlargement of *ci^{rep}* tester wings. (D) Ubiquitous expression of *CG14542^{RNAi}* leads to allometrical overgrowth of wild-type wings, and furthermore to blistering and broadened veins at higher expressivity (E). (F) Enlargement of the eye upon *CG14542^{RNAi}*. (G) Broadened L3 vein (upper inset) and (H) loss of chemosensory bristles (upper inset) upon *CG14542^{RNAi}*. (I,I') Campaniform sensillae are supernumerous along LV3 (I) and ectopic along the anterior wing margin (I'). (J) *CG14542^{RNAi}* expression in the entire wing primordium results in a strongly overgrown L3 wing disc. (K) *CG14542^{RNAi}* expression in wing clones causes pyknotic cells. (K,K') Close up of a wing pouch clone shows caspase activation in *CG14542^{RNAi}* expressing cells. (L) Clones round up and overgrow outside of the wing pouch, whereas clones grow normally within and outside the wing pouch in control wing discs (M). (N) Upregulation of pMad and (O) Wg in *CG14542^{RNAi}* expressing clones. (P) Phospho-Histone expression is enhanced upon *CG14542^{RNAi}* expression in the entire wing primordium, as compared to

control wing discs of the same stage (P'). (Q) Expression of Wg is expanded in the inner ring of the wing primordium. (Q'') Wg expression in a wild-type wing primordium. (R) Overexpression of CG14542 cds causes a likely dominant-negative effect, resulting in wing notches and slightly thickened veins.

Supplementary Figures

Figure S1. Efficient knockdown of known genes required for wing growth
(A) Wild-type wing. (B) *ci*^{act} tester wing. (B') *ci*^{rep} tester wing. (C-F') Knockdown of genes required for Dpp signaling like *mad* (C,C'), *med* (D,D'), *put* (E,E') and *tkv* (F,F') cause wing size reduction of both tester conditions. (G-J) RNAi of known tumor suppressor-like genes like *brk* (G,G'), *warts* (H,H'), *ex* (I,I') lead to enlarged wing sizes in both tester conditions. (J) *hyd*^{RNAi} causes strong blistering in the *ci*^{act} condition.

Figure S2. RNAi of pro-apoptotic genes
(A) Wild-type wing. (B) Expression of activated caspase in a wild-type wing primordium. (B) *ci*^{rep} tester wing. (B') Caspase activation is strongly enhanced in a *ci*^{rep} tester wing primordium. (C) *hid*^{RNAi} or (D) *rpr*^{RNAi} can partially restore wing size. (C') Caspase activation is suppressed in imaginal discs upon expression of *hid*^{RNAi}. (B-D) *nub*-Gal4 was used for transgene expression.

Figure S3. Enrichment/depletion analysis of wing RNAi screen hits
Enrichment/depletion of (A) all hits and phenotypes and (B) phenotype-subclasses "larger" and "smaller".

Figure S4. *ric8a* is required for proper dorso-ventral wing adhesion
(A) Wild-type wing. (A') *ci*^{act} tester wing. (B,B') The primary screen phenotype of *ric8a* knockdown in the wing primordium. (C-F') Knockdown of *ric8a* by three independent oligo-based RNAi transgenes. (E,E') A mutant oligo does not cause a phenotype. (G,H) Expression of *ric8a*^{RNAi} in the dorsal wing primordium suppresses the Myc-dependent down-bending of the dorsal wing blade, and even reduces wing size, but enhances the failure of the dorsal and ventral wing blade apposition. (H') Accumulation of cells between the dorsal and ventral wing blade.

Figure S5. Comparison of the wing RNAi screen with cell culture RNAi screens
Blue circle: all hits from our wing RNAi screen. Red circle: the cell culture RNAi screen family „cell cycle/cell growth" (including hits of 9 RNAi screens). Green circle: screen family „signaling pathways" (including hits of 10 RNAi screens).

Figure S6. Supplementary data to CG14542
Chemosensory bristles normally associated with L1 at the dorsal wing blade margin were lost upon RNAi (Fig. 7H). Three campaniform sensillae which normally form along L3 appear supernumary along L3 and L1 (Fig. 7I,I'). *CG14542*^{RNAi} expression in the notum also led to loss of thoracic sensory bristles (not shown).
(A) pMad expression is upregulated and ectopically activated in *CG14542*^{RNAi} expressing clones. (B) pMad expression domains are distorted in a *CG14542*^{RNAi} expressing wing primordium.

Figure S7. Protein interactions of Ric8a, CG15652 and CG6854
Protein networks are based on STRING.

Figure S_. *CG11286* is required for Notch/Wg signaling
(A) Wild-type wing. (B) *ci*^{act} tester wing. (C) *ci*^{rep} tester wing. (D-F) *CG11296*^{RNAi} causes notched wings and broadened veins in the wild-type and the tester wings. (G-I) The wing margin notch phenotype is strongly enhanced in a sal::*lgs*^{17E} background. (J,J') Expression of the Notch signaling target genes *m8* (J) and *cut* (J') is lost upon expression of *CG11296*^{RNAi} in the posterior wing primordium by hh-Gal4, whereas gene expression in the anterior (control) compartment devoid of RNAi is not affected. Such wing primordia result in adult wings where a comprehensive notch removed most of the posterior compartment (K). (L-M'') Expression of Cut (L-L'') and Sens (M-M'') require *CG11286* in a cell-autonomous way.

Supplementary Tables

Table S1. RNAi wing phenotypes of known components and signaling pathways

Table S2. Genes and their wing phenotypes in the two tester conditions
A color code at the end of each list denotes the strength of the phenotype.
(A) Genes that scored in the S class (tester wing size smaller)

(B) Genes that scored in the L class (tester wing size enlarged)
(C) Genes that scored in the R class (tester wing rounded)
(D) Genes that a bivalent RNAi phenotype (both tester systems suppressed)
(E) Genes of the ambiguous class (different RNAi lines of a gene led to opposite wing size phenotypes)
(F) Genes that scored in the M class (tester wing sizes are not altered, but display an altered morphology, for example blistered wing epithelia)
(G) Genes that scored in the death class, including the lethal class (larval or pupal death in both tester conditions), the partial lethal class (larval or pupal death in only one tester condition), and the necrotic class (predominant necrosis in both tester wing conditions)

Table S3. Wing RNAi hits with human disease orthologs

Table S4. GOstat enrichments (biological process, molecular function)
Comparison of the wing RNAi gene set with the hits of the cell culture RNAi screen family „cell cycle/cell growth“ (including hits of 9 RNAi screens).

Material and Methods

In vivo knockdown of gene function by RNAi

The library of transgenic UAS-RNAi fly lines was provided by the Vienna Drosophila RNAi Center (VDRC, Vienna; www.vdrc.at; Dietzl et al., 2007). „Hairpin RNA“ expression was based on the Gal4/UAS system (Brand and Perrimon, 1993). 21'206 VDRC RNAi lines were available to carry out our screen, which we could map to 11.466 genes (Flybase release ___ of the *Drosophila melanogaster* genome), including ___ "manually mapped" genes which could not be mapped by computational means. At present there are 14.306 putative protein-coding genes in *Drosophila melanogaster* (i) FlyBase annotated: 14'141 genes and ii) 165 „hand-mapped“ genes, which have only an FBgn identifier, but no CG number and no cytological map position.

Tester fly lines for the modifier screen: two conditions for RNA interference

For the screen we generated two genetically modified tester strains that are based on Hh-signaling either below or above the normal level (Figure 1). Genotype of the tester fly lines: (i) small wing tester: *nub-Gal4*/CyO^*UAS-ci*^{rep}/Tm6b; (ii) large wing tester: *nub-Gal4*, *UAS-ci*^{act}/CyO. Sensitivity of cells to levels of Hh, and consequently Dpp, is reflected by the extent of the intervein regions, the vein pattern and by the wing growth. To create two opposing levels of Hh-signaling, we manipulated endogenous functions of Ci (Cubitus interruptus), the transcriptional effector of Hh signaling (Slusarski et al., 1995), by misexpression of Ci-transgenes in wild-type wing primordia. The overexpression of Ci^{cell} (= Ci^{rep}), a dominant repressor form of Ci (Methot and Basler, 1999), prevents Hh signaling. Consequently, Dpp signaling is impaired, which eventually results in small-wing phenotypes with anterior vein fusions and partial vein loss. The expression of Ci^{pka4} in the other tester strain (= Ci^{act}) results in constitutive transcriptional activation of Hh target genes including *dpp* (Methot and Basler, 2000). Consequently, Dpp signaling is hyperactivated, causing a large-wing phenotype. Here, wing veins are mostly affected in the anterior compartment, with enlargement of intervein regions L2/L3. In summary, due to moderate Ci-transgene activities each wing condition lent itself to size-modification into both directions, i.e. size enhancement or reduction. A pilot screen preceding the actual large-scale screen tested 160 UAS-RNAi lines in combination with several wing- and semi-wing-specific Gal4-driver lines, including *scalloped*-Gal4, *MS1096*-Gal4 or *vestigial*-Gal4 (BDSC, Bloomington *Drosophila* Stock Center). We eventually chose *nubbin*-Gal4 (a gift from the lab of G. Morata; Calleja et al., 1996), since it combined appropriate activity, specificity and less general lethality than other Gal4 sources.

Fly crossing scheme for the RNAi screen

2-3 female virgin flies of each of the two tester strains were crossed to 1-2 males per UAS-RNAi fly line (21'206 RNAi fly lines in total) and kept at 26°C. Standard size plastic food tubes with yeast and nipagin were used for the fly crosses.

Wing phenotype analysis upon RNAi knockdown

14 days after setting up the crosses, the F1 generation was analyzed for phenotypical alterations of the wings with a standard stereomicroscope. RNAi based alterations of tester wing phenotypes (i.e., wing size

or morphology) were compared with unaltered tester wing phenotypes and scored in a semiquantitative and descriptive way (Table S2). Adult wings were flat mounted in Euparal and were photographed with a Zeiss camera adjusted to a Zeiss microscope (Zeiss AxioCam, Zeiss Axioplan). If not noted otherwise, flat mounted adult wings were 5x magnified. Tester phenotype controls during the screen: crosses of heterozygous RNAi lines or X-linked RNAi lines allowed us to analyze also flies with tester-only genotypes as internal controls within the same F1 generation. In addition, tester-only control fly crosses were included in each batch of experimental crosses of the same date.

Molecular characterization of RNAi lines

Genomic DNA of VDRC RNAi flies were tested for the presence of their predicted hairpin-insert by standard DNA sequencing analysis of the PCR product obtained after PCR amplification of the inverted repeat sequence. Genomic DNA was obtained from one male fly of the respective RNAi fly line based on a standard protocol of the Berkeley Drosophila Genome Project website (www.fruitfly.org).

Secondary confirmation of the primary RNAi phenotype

For genes which showed a wing phenotype in the primary RNAi screen we used transgenic RNAi lines from NIG-Fly (www.shigen.nig.ac.jp/fly/nigfly) or custom made UAS-RNAi lines based on the oligonucleotide method of Haley et al., 2008.

Quantitative RT-PCR

To test whether gene expression is downregulated upon UAS-hairpin expression, total RNA was isolated from third larval instar wing imaginal discs (Nucleospin RNA II kit, Machery-Nagel) which expressed a particular UAS-RNAi transgene under the control of *c765-Gal4* at 29°C (*c765-Gal4* is active throughout the imaginal disc; Nellen et al., 1996). Subsequent cDNA synthesis was performed by the Transcriptor High Fidelity cDNA Synthesis Kit (Roche). Quantitative real time PCR reactions were performed in triplicates and analyzed by the SYBR Green kit and the ABI Prism 7900HT System of Applied Biosystems. Gene expression of *actin-5c, TBP, gapdh-*α and *cyp1* was used as standards to normalize RNA expression levels in the RNAi conditions.

Inverse PCR (i-PCR)

To localize the genomic integration site of the hairpin transgene of certain RNAi lines i-PCR was carried out according to a standard protocol of the Berkeley Drosophila Genome Project website (www.fruitfly.org). As a note of caution we like to mention RNAi results which are most likely based on endogenous gene activation by nearby inserted hairpin transgenes: (i) we scored the RNAi phenotype of the VDRC line 40852 (*gr23a*) as a Hedgehog signaling gain-of-function phenotype, which strongly resembled *moonrat*, an allele causing ectopic activation of Hh signaling (Felsenfeld and Kennison, 1995). We localized the *gr23a* hairpin transgene in the upstream regulatory region of *hh*, thereby indicating the likely cause of *hh* overexpression (several new but phenotypically silent insertions of the same hairpin construct corroborate this finding); (ii) we scored the RNAi phenotype of the VDRC line 26508 (CG4674) as a Notch signaling loss-of-function-like phenotype (including missing tissue along the wing margin and thickened, terminally delta-like broadened veins). iPCR revealed the integration of the hairpin element upstream of the *fringe* coding region.

CDS-overexpression in flies

For gain-of-function studies coding regions of respective genes were overexpressed based on the site specific ΦC31-integration system (Bischof et al., 2007).

Transgene induction and RNAi phenotype analysis

To analyze reporter gene activities (*lacZ*/bGal) of third instar larval wing imaginal discs we crossed UAS-RNAi males to virgin females of the following genotypes:

m8-lacZ/CyO^hh-Gal4/Tm6b (*m8-lacZ*: gift from G. Halder)

brk^{X47}*-lacZ*; ap-Gal4/CyO (*brk*^{X47}*-lacZ*: Campbell and Tomlinson, 1999)

wg-lacZ/CyO (Kassis, 1990)

For clonal analyses UAS-RNAi transgenes were crossed to virgin females of the genotype *yw hsflp; +/+; act5c>CD2>Gal4, UAS-GFP*. Larvae from this cross were heat-shocked 48 hrs AEL at 37°C for 15 min.

Immunohistochemistry and antibodies

Immunostaining of late third instar larval wing discs were performed as per a standard protocol. The following primary antibodies were used: mouse anti-bGal (1:2000, Promega), guinea pig anti-Senseless (1:800, gift from H. Bellen), mouse anti-Wingless (1:1000, DSHB), rabbit anti-phospho Histone 3 (1:400, Upstate), mouse anti-Cut (1:20, gift from S. Bray), rabbit anti-cleaved Caspase3 (1:400, Cell Signaling), rabbit anti-Spalt (1:50, gift from), rabbit anti phosho-Mad (1:800, gift from E. Laufer). Secondary antibodies (Molecular Probes, 1:400): goat antibodies Alexa Fluor 594 and 499.

Bioinformatic analyses

Definition of off targets: VDRC RNAi Lines which scored positive in the wing growth screen were analyzed with respect to their putative off targets based on the VDRC web page. To be as stringent as possible, while not needlessly eliminating real growth regulators, we defined an RNA line as being off-target prone based on the following criteria: s19 ≤ 0.96 AND/OR CAN ≥ 5 AND/OR on-target.

GO term enrichment analyses:

Gene Ontology (GO) enrichment analysis is based on GoStat with a Benjamini and Hochberg correction using the online GoToolbox application. $p < 0.05$ was considered to indicate statistical significance.

Prediction of orthologous genes:

Orthologs were retrieved from InParanoid.

Figure 1 The experimental set-up of the growth RNAi screen

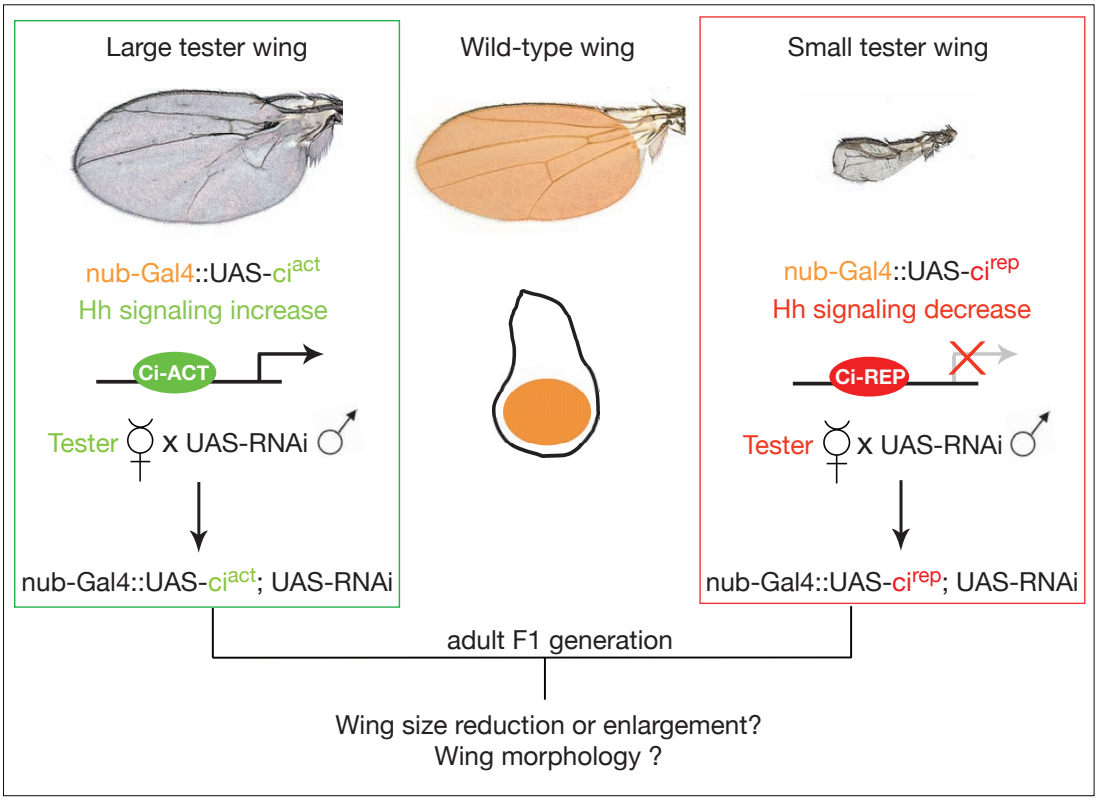


Figure 2

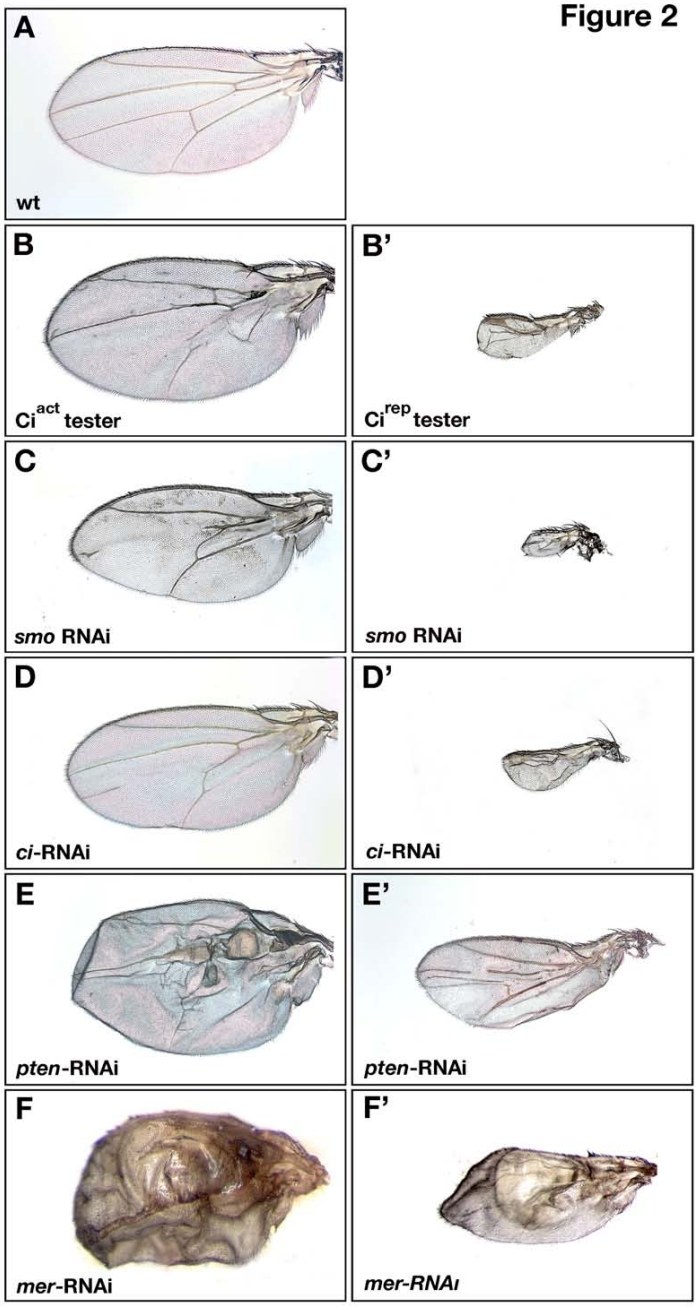

A
wt
B
Ci^act tester
B'
Ci^rep tester
C
smo RNAi
C'
smo RNAi
D
ci-RNAi
D'
ci-RNAi
E
pten-RNAi
E'
pten-RNAi
F
mer-RNAi
F'
mer-RNAi

Figure 3

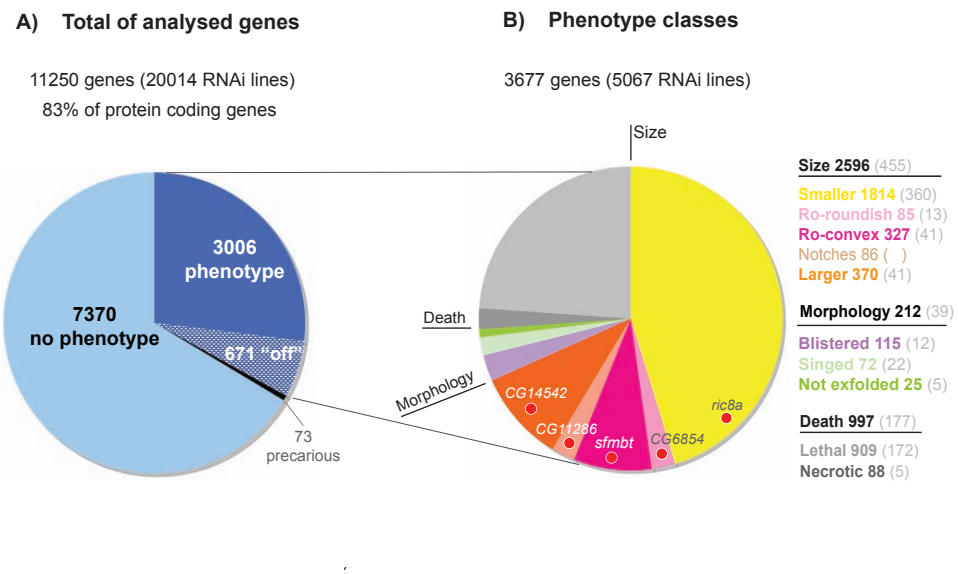

A) Total of analysed genes
11250 genes (20014 RNAi lines)
83% of protein coding genes
3006 phenotype
7370 no phenotype
671 "off"
73 precarious
B) Phenotype classes
3677 genes (5067 RNAi lines)
Size
Death
Morphology
CG14542
CG11286
sfmbt
CG6854
ric8a
Size 2596 (455)
Smaller 1814 (360)
Ro-roundish 85 (13)
Ro-convex 327 (41)
Notches 86 ( )
Larger 370 (41)
Morphology 212 (39)
Blistered 115 (12)
Singed 72 (22)
Not exfolded 25 (5)
Death 997 (177)
Lethal 909 (172)
Necrotic 88 (5)

Figure 4 Phenotype classes

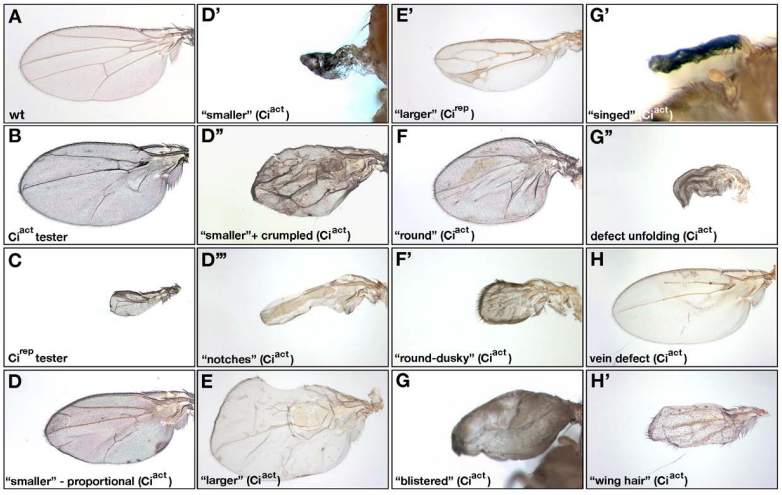


Figure 5

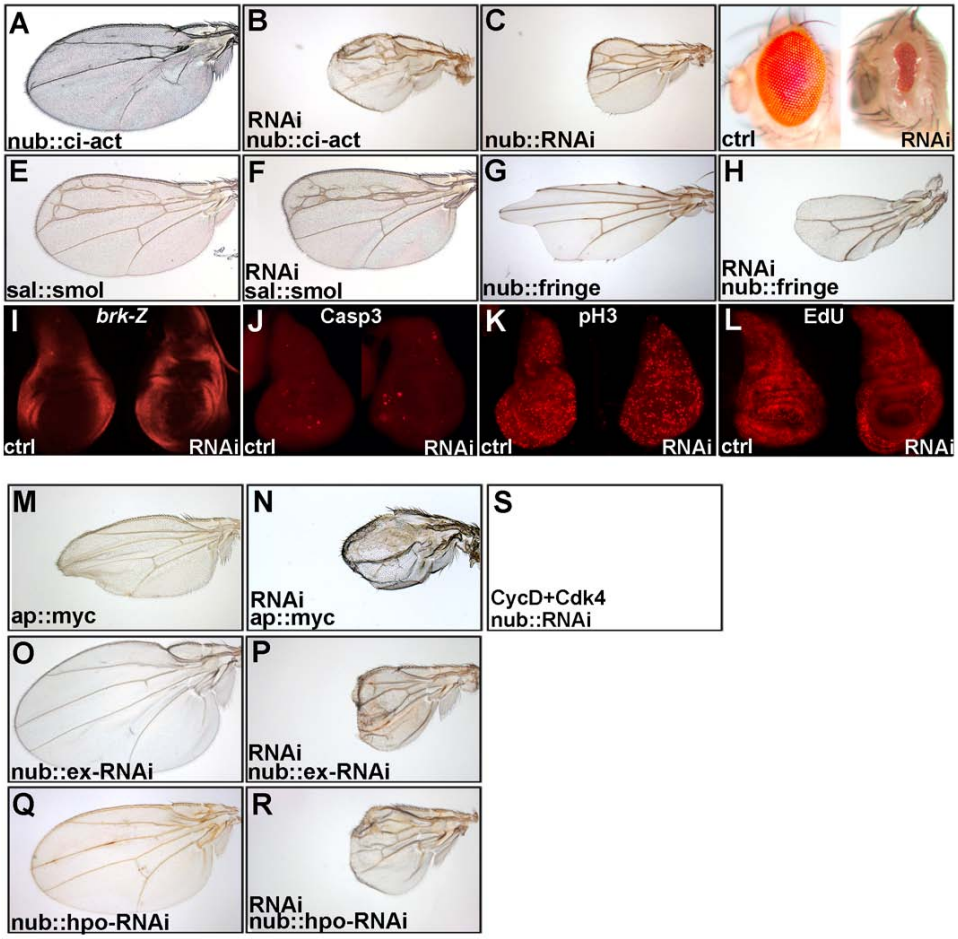

A
nub::ci-act
B
RNAi
nub::ci-act
C
nub::RNAi
ctrl
RNAi
E
sal::smol
F
RNAi
sal::smol
G
nub::fringe
H
RNAi
nub::fringe
I
brk-Z
ctrl
RNAi
J
Casp3
ctrl
RNAi
K
pH3
ctrl
RNAi
L
EdU
ctrl
RNAi
M
ap::myc
N
RNAi
ap::myc
S
CycD+Cdk4
nub::RNAi
O
nub::ex-RNAi
P
RNAi
nub::ex-RNAi
Q
nub::hpo-RNAi
R
RNAi
nub::hpo-RNAi

Figure 7

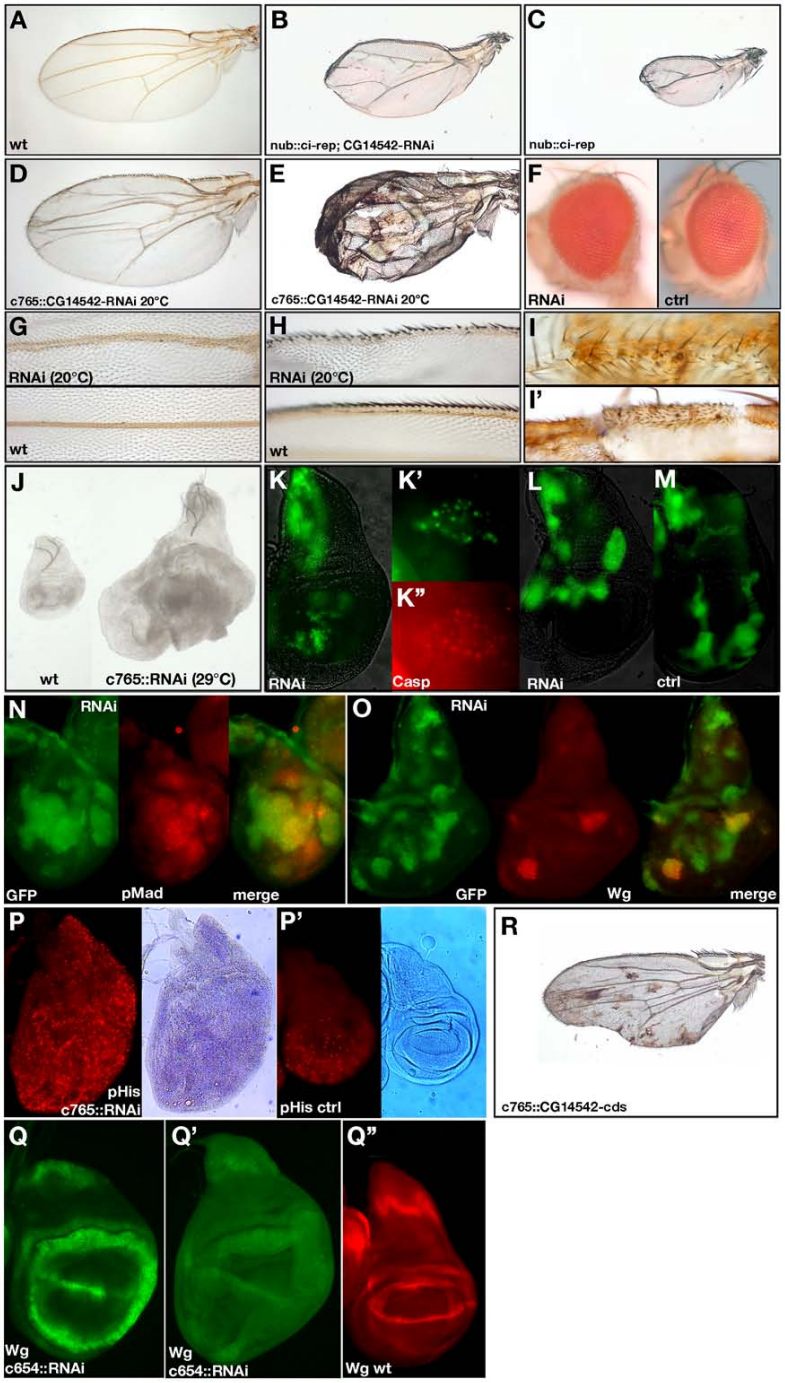

A
wt
B
nub::ci-rep; CG14542-RNAi
C
nub::ci-rep
D
c765::CG14542-RNAi 20°C
E
c765::CG14542-RNAi 20°C
F
RNAi
ctrl
G
RNAi (20°C)
wt
H
RNAi (20°C)
wt
I
I'
J
wt
c765::RNAi (29°C)
K
K'
K"
RNAi
Casp
L
M
RNAi
ctrl
N
RNAi
GFP
pMad
merge
O
RNAi
GFP
Wg
merge
P
pHis
c765::RNAi
P'
pHis ctrl
R
c765::CG14542-cds
Q
Wg
c654::RNAi
Q'
Wg
c654::RNAi
Q"
Wg wt

Figure S1

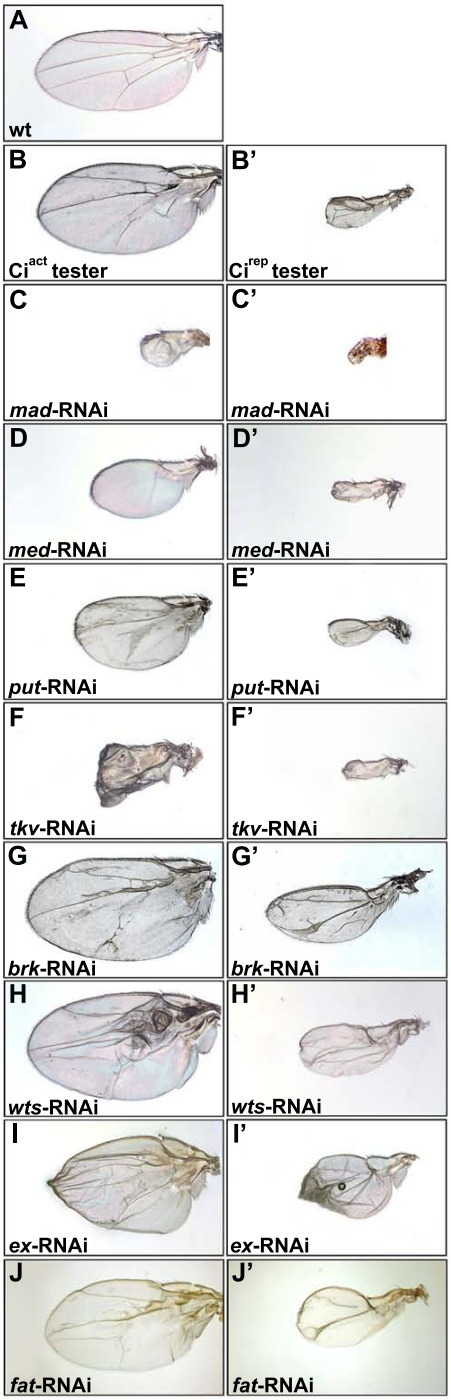
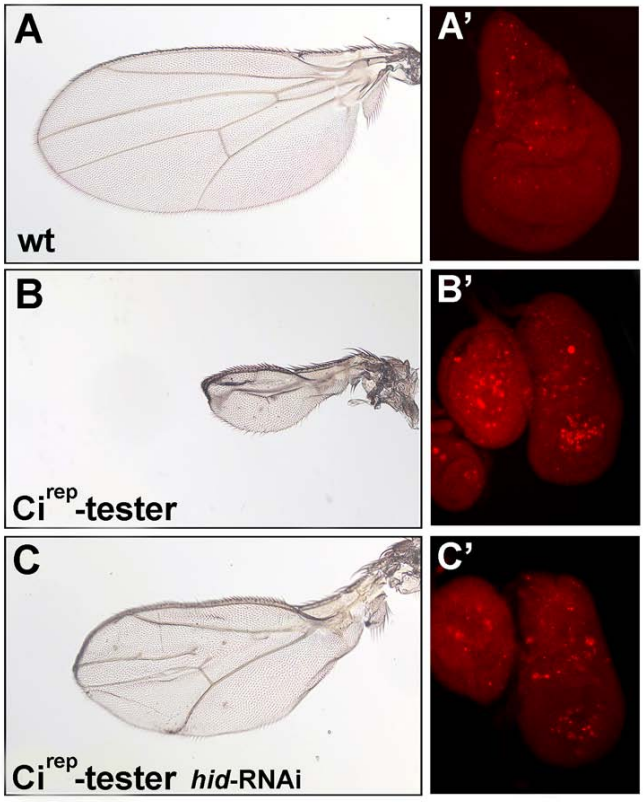
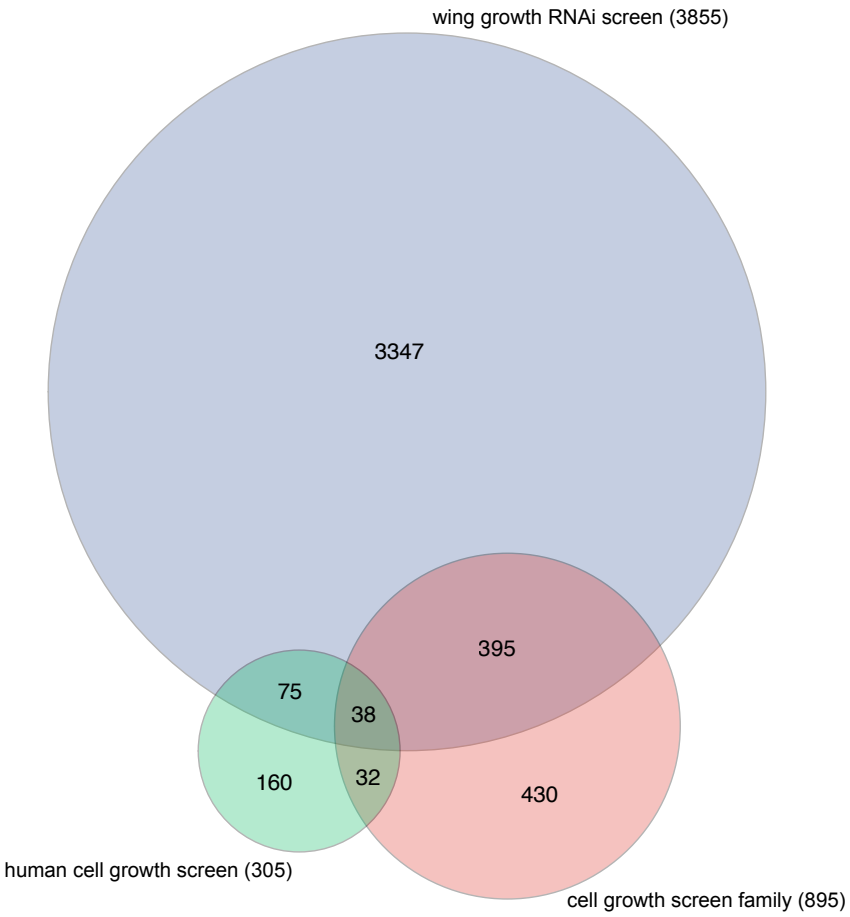

A
wt
B
Ci^act tester
B'
Ci^rep tester
C
mad-RNAi
C'
mad-RNAi
D
med-RNAi
D'
med-RNAi
E
put-RNAi
E'
put-RNAi
F
tkv-RNAi
F'
tkv-RNAi
G
brk-RNAi
G'
brk-RNAi
H
wts-RNAi
H'
wts-RNAi
I
ex-RNAi
I'
ex-RNAi
J
fat-RNAi
J'
fat-RNAi

Figure S2




wing growth RNAi screen (3855)
3347
395
75
38
160
32
430
human cell growth screen (305)
cell growth screen family (895)

SMALLER_GO enrichment_Biological Process_overrepresented

	p-value	Enrichment value	
GO:0007448	0.005414623	5.690703736	anterior/posterior pattern formation, imaginal disc
GO:0007478	0.030916795	5.685217391	leg disc morphogenesis
GO:0035156	0.031183016	5.660914582	fusion cell fate specification
GO:0007228	0.005524144	5.652885444	positive regulation of hh target transcription factor activity
GO:0045880	0.005524144	5.652885444	positive regulation of smoothened signaling pathway
GO:0008284	0.005524144	5.652885444	positive regulation of cell proliferation
GO:0035217	0.005524144	5.652885444	labial disc development
GO:0048099	0.005524144	5.652885444	anterior/posterior lineage restriction, imaginal disc
GO:0046672	0.005524144	5.652885444	positive regulation of compound eye retinal cell programmed ce
GO:0021954	0.005524144	5.652885444	central nervous system neuron development
GO:0007386	0.005524144	5.652885444	compartment specification
GO:0035019	0.031271706	5.652885444	somatic stem cell maintenance
GO:0045850	0.031271706	5.652885444	positive regulation of nurse cell apoptosis
GO:0016340	0.031271706	5.652885444	calcium-dependent cell-matrix adhesion
GO:0009186	0.031271706	5.652885444	deoxyribonucleoside diphosphate metabolic process
GO:0035191	0.031271706	5.652885444	nuclear axial expansion
GO:0035157	0.031271706	5.652885444	negative regulation of fusion cell fate specification
GO:0007032	0.031271706	5.652885444	endosome organization and biogenesis
GO:0042069	0.031271706	5.652885444	regulation of catecholamine metabolic process
GO:0035207	0.031271706	5.652885444	negative regulation of hemocyte proliferation
GO:0021960	0.031271706	5.652885444	anterior commissure morphogenesis
GO:0048100	1.39E-06	5.087596899	wing disc anterior/posterior pattern formation
GO:0030708	0.000180205	4.84533038	germarium-derived female germ-line cyst encapsulation
GO:0007487	0.00088048	4.71073787	analia development
GO:0051090	0.00088048	4.71073787	regulation of transcription factor activity
GO:0045887	0.004187672	4.522308355	positive regulation of synaptic growth at neuromuscular junctior
GO:0009608	0.004187672	4.522308355	response to symbiont
GO:0046622	0.004187672	4.522308355	positive regulation of organ growth
GO:0001746	0.000612376	4.239664083	Bolwig's organ morphogenesis
GO:0035171	0.000612376	4.239664083	lamellocyte differentiation
GO:0015014	0.019171137	4.239664083	heparan sulfate proteoglycan biosynthetic process, polysaccha
GO:0035263	0.019171137	4.239664083	genital disc sexually dimorphic development
GO:0016572	0.019171137	4.239664083	histone phosphorylation
GO:0048076	0.019171137	4.239664083	regulation of compound eye pigmentation
GO:0007483	0.019171137	4.239664083	genital disc morphogenesis
GO:0007418	0.019171137	4.239664083	ventral midline development
GO:0007076	3.46E-06	4.145449325	mitotic chromosome condensation
GO:0030307	9.33E-05	4.111189414	positive regulation of cell growth
GO:0045705	0.002631168	4.037775317	negative regulation of salivary gland boundary specification
GO:0045500	0.002631168	4.037775317	sevenless signaling pathway = EGFR signaling pathway
GO:0035224	0.002631168	4.037775317	genital disc anterior/posterior pattern formation
GO:0007425	0.002631168	4.037775317	epithelial cell fate determination, open tracheal system
GO:0008362	0.000388644	3.957019811	chitin-based embryonic cuticle biosynthetic process
GO:0008101	0.000388644	3.957019811	decapentaplegic receptor signaling pathway
GO:0030713	5.95E-05	3.913536076	ovarian follicle cell stalk formation
GO:0048190	0.000236331	3.768590296	wing disc dorsal/ventral pattern formation
GO:0007307	0.000236331	3.768590296	eggshell chorion gene amplification
GO:0045747	0.010802056	3.768590296	positive regulation of Notch signaling pathway
GO:0046534	0.010802056	3.768590296	positive regulation of photoreceptor cell differentiation
GO:0007458	0.010802056	3.768590296	progression of morphogenetic furrow during compound eye mo
GO:0006379	0.000905567	3.597290737	mRNA cleavage
GO:0035110	0.005843861	3.55326087	leg morphogenesis
GO:0035288	0.005995823	3.533053402	anterior head segmentation
GO:0035214	0.038605768	3.488080073	eye-antennal disc development
GO:0048863	2.64E-06	3.391731266	stem cell differentiation
GO:0030718	2.76E-05	3.391731266	germ-line stem cell maintenance
GO:0035080	0.041646334	3.391731266	heat shock-mediated polytene chromosome puffing
GO:0007523	0.041646334	3.391731266	larval visceral muscle development
GO:0045498	0.041646334	3.391731266	sex comb development
GO:0007473	0.041646334	3.391731266	wing disc proximal/distal pattern formation
GO:0051642	0.041646334	3.391731266	centrosome localization
GO:0007157	0.001842497	3.297516509	heterophilic cell adhesion
GO:0030720	0.001842497	3.297516509	oocyte localization during germarium-derived egg chamber forn
GO:0051297	0.02142646	3.241729649	centrosome organization and biogenesis
GO:0046328	0.011364542	3.151515152	regulation of JNK cascade
GO:0007549	0.011538391	3.140491913	dosage compensation
GO:0048070	0.011538391	3.140491913	regulation of pigmentation during development

Table S1 RNAi wing phenotypes of control genes

Hh signaling (24 genes addressed by VDRC RNAi lines)

Gene symbol	CG	Line	ci-rep	ci-act
boi	CG7894	42248	<	0
boi	CG7894	29592	0	0
boi	CG7894	29593	0	0
boi	CG7894	869	0	0
boi	CG7894	3060	0	0
botv	CG15110	37185	0	<-<< and PD
botv	CG15110	37186	0	<< and PD
ci	CG2125	51479	0	<<
Ckl alpha	CG2028	13664	PD	nd
dally	CG4974	14136	0	<
dbr	CG11317	9293	<<	<<
dbr	CG11317	9292	0	0
dbr	CG11317	18070	<	<<
dbr	CG11317	18071	0	<
disp	CG2019	10004	nd	0
dlp	CG32146	10298	<	<
dlp	CG32146	10299	0	0
fu	CG6551	27662	<<, nec tips	<
fu	CG6551	27663	<	<
hh	CG4637	1402	PD	<
hh	CG4637	43255	<	<
hh	CG4637	1403	PD	PD
iHog	CG9211	1001	<<	PD
iHog	CG9211	29897	< (variable)	0
iHog	CG9211	29898	0	0
nejire	CG15319	3787	<	<<<, most PD
Pka-C1	CG4379	6993	PD	PD
skpA	CG16983	46605	<<	<<<, cr, nec
skpA	CG16983	46607	<<	PD
skpA	CG16983	32789	PD	PD
skpA	CG16983	32790	PD	PD
slmb	CG3412	34273	<	<
slmb	CG3412	34274	<	<
smo	CG11561	9542	<	<
tws	CG6235	34340	0	<, cr
ski	CG11495	6459	0	0
Su(fu)	CG6054	35055	0	0
mts	CG7109	41924	<<	<<
mts	CG7109	35171	0	0
mts	CG7109	35172	0	0
smw	CG10084	38334	<<	<<<
smw	CG10084	38336	<	<
roadkill	CG12537	45759	<<<	nd
roadkill	CG12537	45761	<<(<)	<<
roadkill	CG12537	28800	<<	PD
roadkill	CG12537	28798	<<	PD
cullin 3	CG11861	16331	0	PD
combgap	CG8367	2965	PD	PD
combgap	CG8367	2966	LD	LD

Phenotype codes:

< - <<<	wing size reduction
>->>	wing size enlargement
PD	pupal death
nec	necrosis
bli	wing blisters
cr	crinkled wings
round	round shaped wing
0	no alteration of tester wing
nd	not determined

16 genes	expected phenotype
1 gene	false positive *(slimb)*
3 genes	false negative *(disp, ski, Su(fu))*
4 genes	larval or pupal lethality

Dpp signaling (15 genes addressed by VDRC RNAi lines)

Gene symbol	CG	Line	ci-rep	ci-act
brk	CG9653	2919	>>	>
dad	CG5201	42840	<<<	<<<
dally	CG4974	14136	0	<
dlp	CG32146	10298	<	<
dlp	CG32146	10299	0	0
hth	CG17117	12763	0	>, round
hth	CG17117	12764	0	> and PD
lack, smurf	CG4943	24681	>	>
lack, smurf	CG4943	24680	0	0
mad	CG12399	12635	<<<	<<<
med	CG1775	19688	<	round
med	CG1775	19689	<<	<<
put	CG7904	848	nd	<, nec
put	CG7904	849	<	<<
put	CG7904	37279	0	<<
sax	CG1891	9434	<	<
sax	CG1891	42456	<	<
sax	CG1891	9433	0	bli
mtv	CG5580	41845	0	round, >
shn	CG7734	3226	<	0
tkv	CG14026	862	<	<<, nec, bli
tkv	CG14026	3059	0	<, round
ttv	CG10117	4871	<<	<
ubx	CG10388	37823	haltere >>	0
ubx	CG10388	37825	haltere >	0

14 genes	expected phenotype
1 gene	false positive

Hippo signaling (10 genes addressed by VDRC RNAi lines)

Gene symbol	CG	Line	ci-rep	ci-act
ex	CG4114	22994	>>	>, bli
fat	CG3352	9396	>	PD; bli
dachsous	CG17941	4312	nd	nd
dachsous	CG17941	4313	>	bli
dachsous	CG17941	36219	>	PD; bli
sav	CG33193	46445	bli	bli, rond
sav	CG33193	24801	0	round
sav	CG33193	24802	>	nec spindle
sav	CG33193	46446	0	0
wts	CG12072	9928	>	>>, bli
dachs	CG31610	12555	<<	0
dachs	CG31610	12556	<<	<
fj	CG10917	6774	0	0
mer	CG14228	7161	>	>
mer	CG14228	7162	0	PD
hippo	CG11228	7823	0	0
yorkie	CG4005	40497	<<<	nd

7 genes	expected phenotype
1 genes	false positive
2 genes	false negative *(hippo, fj)*

Total genes (excl. 4 lethal cases): 43 genes
Expected phenotype: 35 genes (81.4%)
False positive: 3 genes (7%)
False negative: 5 genes (11.6%)

Total lines (excl. lines of 4 lethal cases): 80 lines
Expected phenotype: 60 lines (75%)
False positive: 4 lines (5%)
False negative: 16 lines (20%)

Table S2 a) SMALLER class

1905 Genes (2277 Lines): wing size reduction in at least one tester condition

Genes are listed 1.by CG number > 2.by TF-ID number

CG	Gene Symbol	RNAi Line	ci-rep	ci-act
CG10001	AR-2	1327	0	<
CG10002	fkh	37063	<<	PD
CG10009	Noa36	17954	<<<	PD
CG10009	Noa36	19485	<<<	<<, dus, cr, PD
CG10033	for	38319	<<	<<, dus
CG10033	for	38320	<	<, dus
CG10035	CG10035	7524	0	<
CG10037	vvl	47182	<<	PD
CG10052	Rx	44716	nd	<
CG10055	CG10055	17973	<<	PD
CG10064	CG10064	38322	0	<<
CG10084	swm	38334	<<	<<<
CG10084	swm	38336	<	<
CG10107	CG10107	18005	<<<	PD
CG10109	L	38339	0	<, wr, bli
CG10110	cpsf	18009	<<<	nd
CG10117	ttv	4871	<<	<
CG10118	ple	3308	PD	<<
CG10122	Rpl1	12688	0	<<<
CG10125	zpg	33277	<	PD
CG10126	CG10126	44103	<	0
CG10144	CG10144	18019	<<	sin
CG10146	AttA	50320	<<	<<<
CG10152	beat-IV	1410	<	PD
CG10154	CG10154	31188	nd	<
CG10158	CG10158	18026	<<	<<, wr
CG10158	CG10158	47389	nd	sin
CG10174	CG10174	31195	<<	<<
CG10200	CG10200	47613	0	<, wr
CG10209	CG10209	28153	PD	sin
CG10209	CG10209	29042	<	nd
CG10212	SMC2	10711	<<	nd
CG10222	CG10222	18038	<	sin
CG10223	Top2	30625	<<	<<<
CG10226	CG10226	5055	nd	<
CG10228	l(2)k08015	38366	<<	<, wr
CG10230	Rpn9	31206	<<	PD
CG10231	Pde11	18041	<<<	nd
CG1024	CG1024	18599	<<	nd
CG10244	Cad96Ca	1091	0	<<
CG10244	Cad96Ca	49398	0	<
CG10253	CG10253	3321	0	<<
CG10267	CG10267	22836	<<<	nd
CG10272	CG10272	16001	<<<	nd
CG10275	CG10275	36247	<<	PD
CG10279	Rm62	46908	<<	0
CG10281	TfIIFalpha	51209	<, dus	<<<, bli, dus
CG10281	TfIIFalpha	51211	<	<<<, bli, dus
CG10293	how	13756	<<<	nd
CG10311	CG10311	37503	<<	PD
CG10320	CG10320	8837	nd	<<<
CG10324	CG10324	31226	<<<	PD
CG10332	CG10332	49651	<<<	<<<
CG10335	Pbgs	40612	<	nd
CG1034	bcd	12743	0	<
CG10348	CG10348	39663	0	<
CG10354	CG10354	27254	0	<<, wr
CG10374	Lsd-1	30884	<<<	PD
CG10374	Lsd-1	30885	<<	nd
CG10377	Hrb27C	16040	<<<	PD
CG10385	msl-1	9239	<	<
CG10393	amos	11796	<<	nec
CG10414	CG10414	16047	<<<	PD
CG10414	CG10414	47391	0	<, wr, nec
CG10415	TfIIEalpha	12591	nec	<<
CG10449	Catsup	7183	<<<	nd
CG10463	CG10463	31248	0	<
CG10463	CG10463	31249	<<, nec	sin
CG10466	CG10466	23367	<<<	PD
CG10479	CG10479	45098	0	<<
CG1048	zen2	10406	<	nd
CG10480	Bj1	38388	0	<<<
CG10493	CG10493	45363	<<	PD
CG10528	fs(2)ltoPP43	16068	<<	PD
CG10539	S6k	18126	<	<
CG10542	CG10542	15620	nd	<
CG1057	MED31	27284	<	0
CG1057	MED31	27285	nd	<, nec
CG10582	Sin	52094	<<	nd
CG10583	Sse	45091	<	nd
CG10583	Sse	45092	<<	<<, wr
CG10603	mRpL13	31285	0	<, wr, PD
CG1064	Snr1	12645	<<	nd
CG10648	CG10648	51701	<<	PD
CG1065	Scsalpha	27298	<<<	0
CG10688	CG10688	39715	<<	nd
CG10688	CG10688	39716	<<	PD
CG10689	CG10689	31325	<<<	PD
CG10691	l(2)37Cc	12358	<<	PD
CG10701	Moe	37917	<, bif	<<
CG10711	CG10711	16846	<<	nd
CG10724	CG10724	22850	<	PD

Phenotype codes:

<	mild wing size reduction
<<	significant wing size reduction (below wt)
<<<	vestigial wings
bli	blistered, fluid filled, vesicular wings
wr	wrinkled wings
dus	dusky wings (wing hairs aberrant: density a/o length)
sin	singed (dark, spindled) wing
nex	not exfolded wings
N, N*	notches incl. thicker veins (N* slight notches)
bif	bifid or bifurcated wings
nec	predominant necrotic wings
PD (LD)	pupal (or larval) death
0	no alteration of tester phenotype
nd	not determined

Table S3 Positive wing RNAi hits with human disease orthologs

FBgn0004656	CG2252 fs(1)h ENSP00000263377
FBgn0041092	CG13109 tai ENSP00000267974
FBgn0027091	CG8431 Aats-cys ENSP00000369897
FBgn0038827	CG17269 Fancd2 ENSP00000287647
FBgn0086904	CG8759 Nacalpha ENSP00000349212
FBgn0014879	CG4299 Set ENSP00000361777
FBgn0000251	CG1759 cad ENSP00000370408
FBgn0024326	CG9738 Mkk4 ENSP00000262445
FBgn0003444	CG11561 smo ENSP00000249373
FBgn0039120	CG10198 Nup98 ENSP00000316032
FBgn0011829	CG14396 Ret ENSP00000344798
FBgn0026317	CG6147 Tsc1 ENSP00000298552
FBgn0032497	CG6043 CG6043 ENSP00000327563
FBgn0011661	CG10701 Moe ENSP00000353408
FBgn0086384	CG14228 Mer ENSP00000335652
FBgn0034802	CG3800 CG3800 ENSP00000303844
FBgn0034585	CG4030 CG4030 ENSP00000262477
FBgn0004227	CG4211 nonA ENSP00000349748
FBgn0000319	CG9012 Chc ENSP00000269122
FBgn0030855	CG5800 CG5800 ENSP00000314348
FBgn0020245	CG10117 ttv ENSP00000367446
FBgn0015624	CG15319 nej ENSP00000263253
FBgn0000117	CG11579 arm ENSP00000344456
FBgn0015031	CG14028 cype ENSP00000297564
FBgn0005198	CG6975 gig ENSP00000219476
FBgn0038197	CG3143 foxo ENSP00000339527
FBgn0005658	CG7018 Ets65A ENSP00000295727
FBgn0004647	CG3936 N ENSP00000277541
FBgn0003205	CG9375 Ras85D ENSP00000256078
FBgn0004924	CG6146 Top1 ENSP00000354522
FBgn0010389	CG7223 htl ENSP00000309878,ENSP00000311337
FBgn0004419	CG4916 me31B ENSP00000264018
FBgn0023097	CG5206 bon ENSP00000340507
FBgn0000242	CG6500 Bx ENSP00000338207
FBgn0001233	CG1242 Hsp83 ENSP00000335153
FBgn0004864	CG1594 hop ENSP00000371067
FBgn0036685	CG6664 CG6664 ENSP00000263102
FBgn0037657	CG11990 hyx ENSP00000356406
FBgn0003079	CG2845 phl ENSP00000288602
FBgn0011655	CG1775 Med ENSP00000341551
FBgn0000472	CG10798 dm ENSP00000367207
FBgn0016694	CG17888 Pdp1 ENSP00000226067
FBgn0001942	CG9075 eIF-4a ENSP00000326381
FBgn0026441	CG4913 ear ENSP00000252674,ENSP00000369695
FBgn0034975	CG11290 enok ENSP00000265713
FBgn0033842	CG5970 CG5970 ENSP00000304704
FBgn0001123	CG2835 G-salpha60A ENSP00000302237
FBgn0026379	CG5671 Pten ENSP00000361021
FBgn0039559	CG4976 Mes-4 ENSP00000348031

Table S4 GO stats enrichments (Biological Process, Molecular Function)

wing-only gene set

BP

GO-ID	p-value	GO-Term
GO:0006811	0.00043179	ion transport
GO:0022900	0.00526259	electron transport chain
GO:0044262	0.00592389	cellular carbohydrate metabolic process
GO:0042773	0.00631355	ATP synthesis coupled electron transport
GO:0030001	0.00748698	metal ion transport
GO:0019752	0.01157004	carboxylic acid metabolic process
GO:0006030	0.02253184	chitin metabolic process
GO:0006631	0.02701386	fatty acid metabolic process
GO:0006066	0.03750415	alcohol metabolic process
GO:0006629	0.0378727	lipid metabolic process
GO:0005975	0.03864249	carbohydrate metabolic process
GO:0048066	0.03882044	pigmentation during development
GO:0006839	0.03882044	mitochondrial transport
GO:0019725	0.0415997	cellular homeostasis

MF

GO-ID	p-value	GO-Term
GO:0016491	3.59E-05	oxidoreductase activity
GO:0022891	0.00034535	substrate-specific transmembrane transporter activity
GO:0043169	0.00049788	cation binding
GO:0046872	0.00062499	metal ion binding
GO:0005102	0.00084179	receptor binding
GO:0015267	0.00365837	channel activity
GO:0043565	0.00369695	sequence-specific DNA binding
GO:0003700	0.00404025	transcription factor activity
GO:0005506	0.01058971	iron ion binding
GO:0009055	0.01063731	electron carrier activity
GO:0008270	0.01604291	zinc ion binding
GO:0005215	0.01615857	transporter activity
GO:0022857	0.01964121	transmembrane transporter activity
GO:0005214	0.01989252	structural constituent of chitin-based cuticle
GO:0001584	0.02228613	rhodopsin-like receptor activity
GO:0048037	0.03194851	cofactor binding
GO:0004497	0.03732956	monooxygenase activity
GO:0015276	0.04368422	ligand-gated ion channel activity
GO:0030247	0.04368422	polysaccharide binding
GO:0008324	0.04373319	cation transmembrane transporter activity
GO:0005261	0.04866047	cation channel activity

culture growth screens-only gene set

BP

GO-ID	p-value	GO-Term
GO:0006468	0.00029105	protein amino acid phosphorylation
GO:0006796	0.00141617	phosphate metabolic process
GO:0006464	0.00219547	protein modification process
GO:0048066	0.01199176	pigmentation during development
GO:0046843	0.01411882	dorsal appendage formation
GO:0006888	0.02515345	ER to Golgi vesicle-mediated transport
GO:0051225	0.04156771	spindle assembly
GO:0007143	0.04354042	female meiosis

MF

GO-ID	p-value	GO-Term
GO:0016773	0.00010095	phosphotransferase activity, alcohol group as acceptor
GO:0016301	0.00015296	kinase activity
GO:0016740	0.00070012	transferase activity
GO:0004674	0.00121754	protein serine/threonine kinase activity
GO:0005102	0.00303437	receptor binding
GO:0004683	0.02696999	calmodulin-dependent protein kinase activity
GO:0017076	0.03405156	purine nucleotide binding

B

CHARACTERIZATION OF THE RAPAMYCIN-SENSITIVE PHOSPHOPROTEOME REVEALS THAT SCH9 IS A CENTRAL COORDINATOR OF PROTEIN SYNTHESIS

For this work I processed the LC–MS/MS data generated using the clustering algorithm for LC–MS/MS alignments described in chapter 7.

The original publication is included below.

Characterization of the rapamycin-sensitive phosphoproteome reveals that Sch9 is a central coordinator of protein synthesis

Alexandre Huber,[1,8] Bernd Bodenmiller,[2,8] Aino Uotila,[1] Michael Stahl,[1] Stefanie Wanka,[3] Bertran Gerrits,[4] Ruedi Aebersold,[2,5,6,7] and Robbie Loewith[1,9]

[1]Department of Molecular Biology, University of Geneva, Geneva 1211, Switzerland; [2]Institute of Molecular Systems Biology, ETH Zürich, Zürich 8093, Switzerland; [3]Institute of Molecular Biology, University of Zurich, Zürich 8057, Switzerland; [4]Functional Genomics Center Zurich, University of Zürich, Zürich 8057, Switzerland; [5]Institute for Systems Biology, Seattle, Washington 98103, USA; [6]Competence Center for Systems Physiology and Metabolic Diseases, ETH Zürich, Zürich 8093, Switzerland; [7]Faculty of Science, University of Zürich, Zürich 8057, Switzerland

The target of rapamycin complex 1 (TORC1) is an essential multiprotein complex conserved from yeast to humans. Under favorable growth conditions, and in the absence of the macrolide rapamycin, TORC1 is active, and influences virtually all aspects of cell growth. Although two direct effectors of yeast TORC1 have been reported (Tap42, a regulator of PP2A phosphatases and Sch9, an AGC family kinase), the signaling pathways that couple TORC1 to its distal effectors were not well understood. To elucidate these pathways we developed and employed a quantitative, label-free mass spectrometry approach. Analyses of the rapamycin-sensitive phosphoproteomes in various genetic backgrounds revealed both documented and novel TORC1 effectors and allowed us to partition phosphorylation events between Tap42 and Sch9. Follow-up detailed characterization shows that Sch9 regulates RNA polymerases I and III, the latter via Maf1, in addition to translation initiation and the expression of ribosomal protein and ribosome biogenesis genes. This demonstrates that Sch9 is a master regulator of protein synthesis.

[*Keywords*: Maf1; Sch9; TOR; phosphoproteomics; rapamycin; ribosome biogenesis]

Supplemental material is available at http://www.genesdev.org.

Received March 30, 2009; revised version accepted June 19, 2009.

The target of rapamycin complex 1 (TORC1) and TORC2 are large essential multiprotein assemblies structurally and functionally conserved throughout eukaryotic evolution. At their structural core are the TOR proteins, which are large Ser/Thr kinases belonging to the phosphatidylinositol kinase-like kinase family (Wullschleger et al. 2006). It is generally believed that TORC1 is active when sufficient nutrients are present and noxious stressors are absent (DH Kim et al. 2002; Urban et al. 2007). TORC1 also appears to monitor intracellular cues: Treatment with cycloheximide, a translation elongation inhibitor, causes a potent increase in TORC1 activity (Beugnet et al. 2003; Urban et al. 2007).

Rapamycin has been an invaluable tool to study pathways downstream from TORC1. This hydrophobic macrolide easily crosses cell membranes to rapidly and specifically inhibit TORC1. Application of rapamycin to yeast cells demonstrated that TORC1 signals promote cell growth through both the stimulation of anabolic processes, such as protein synthesis and ribosome biogenesis, and the inhibition of catabolic processes, such as autophagy, and stress-responsive transcription programs. For a comprehensive review on TORC1 signaling in yeast, see De Virgilio and Loewith (2006).

The molecular pathways linking TORC1 to its distal readouts are presently only partially characterized. At least two direct effectors downstream from TORC1 have been described: Tap42 and Sch9 (Di Como and Arndt 1996; Jiang and Broach 1999; Urban et al. 2007). Tap42 binds and regulates PP2A and PP2A-like protein phosphatases (Nanahoshi et al. 1998). Tap42 also interacts with and is directly phosphorylated by TORC1 (Jiang and Broach 1999). Genetic evidence clearly shows that Tap42 mediates TORC1 signals to a number of distal readouts. This is based on the observation that *tap42-11*, a temperature-sensitive allele of *TAP42*, confers semidominant resistance to rapamycin at the permissive temperature of 25°C (Di Como and Arndt 1996; Duvel et al. 2003). Indeed, in *tap42-11*

[8]These authors contributed equally to this work.
[9]Corresponding author.
E-MAIL robbie.loewith@unige.ch; FAX 41-22-379-68-68.
Article is online at http://www.genesdev.org/cgi/doi/10.1101/gad.532109.

cells, the activation of the transcription factors Gcn4, Gln3, Gat1, and Msn2/4, and the kinase Npr1, normally observed after inhibition of TORC1 with rapamycin, is partially or, in some cases, completely blocked (Schmidt et al. 1998; Cherkasova and Hinnebusch 2003; Duvel et al. 2003; Santhanam et al. 2004).

Sch9 is a Ser/Thr protein kinase of the AGC family. It is directly phosphorylated by TORC1 at its C terminus on at least five residues, and these phosphorylation events are required for catalytic activity (Urban et al. 2007). Replacing some or all of these residues with acidic amino acids (*SCH9*^{3E} and *SCH9*^{DE} alleles) yields versions of Sch9 that retain activity even in the absence of upstream signals from TORC1 (Urban et al. 2007). Sch9 mediates TORC1 signals to a number of distal readouts: Sch9 blocks the induction of genes required for entry into G_0 by directly phosphorylating, and thereby antagonizing, the nuclear accumulation of the Ser/Thr kinase Rim15 (Wanke et al. 2008; Wei et al. 2008); Sch9 is critical for TORC1 to antagonize eIF2α phosphorylation and thus maintain efficient translation initiation (Urban et al. 2007); and Sch9 plays important roles in the regulated expression of RNA polymerase II (Pol II)-dependent genes required for ribosome biogenesis (Jorgensen et al. 2004; Urban et al. 2007). Except for Rim15, the substrates of Sch9 involved in these processes are not known.

Ribosome biogenesis is a highly coordinated process requiring the concerted activity of the three nuclear RNA polymerases (RNA Pol I–III) (Planta 1997; Venema and Tollervey 1999). As it consumes a high amount of cellular energy (Warner 1999), it is not surprising that ribosome biogenesis is tightly coupled to environmental growth conditions. Much of this regulation is mediated by TORC1 (Zaragoza et al. 1998; Mayer and Grummt 2006).

RNA Pol I is dedicated to the transcription of 35S pre-rRNA, which is subsequently processed to 25S, 18S, and 5.8S rRNAs (Venema and Tollervey 1999). Among other models, RNA Pol I was proposed to be regulated by TORC1 via recruitment of the essential initiation factor Rrn3 (Claypool et al. 2004).

RNA Pol II transcription is required for expression of *ribosomal protein* (*RP*) genes and *ribosome biogenesis* (*ribi*) genes that encode proteins required for nucleolar rRNA processing and assembly of ribosomal subunits (Jorgensen et al. 2004). TORC1 regulates the expression of both *RP* and *ribi* genes by controlling the activities of several transcription factors. Some of this regulation is mediated by Sch9 (Jorgensen et al. 2004; Urban et al. 2007).

RNA Pol III transcribes the 5S rRNA, tRNAs, and other stable noncoding RNAs (Geiduschek and Kassavetis 2001). TORC1 regulates RNA Pol III via its conserved repressor, Maf1 (Upadhya et al. 2002). Under favorable growth conditions, Maf1 is highly phosphorylated and is shuttled out of the nucleus. Inactivation of TORC1 results in the rapid dephosphorylation and nuclear accumulation of Maf1 (Oficjalska-Pham et al. 2006; Roberts et al. 2006). TORC1 has been proposed to maintain Maf1 phosphorylation by antagonizing the activity of PP2A family phosphatases, while PKA has been proposed to be the Maf1 kinase (Moir et al. 2006; Oficjalska-Pham et al. 2006). Dephosphorylated Maf1 binds RNA Pol III and, consequently, blocks RNA Pol III transcription via poorly defined mechanisms (Desai et al. 2005; Oficjalska-Pham et al. 2006; Roberts et al. 2006).

Although many distal readouts downstream from TORC1 are known, undoubtedly more remain to be identified. Additionally, the signaling cascades that couple TORC1 to its known readouts remain incompletely understood. To better characterize these pathways we wished to define the TORC1-regulated phosphoproteome. Several liquid chromatography tandem mass spectrometry (LC-MS/MS) approaches have been developed recently to quantify protein phosphorylation (Domon and Aebersold 2006; Olsen et al. 2006). Most prominent among these are protocols based on the differential isotopic labeling of phosphopeptides enriched from protein digests that enables relative quantification between biological samples (Zhou et al. 2001; Olsen et al. 2006). Beyond the restricted number of biological comparisons possible, these methods suffer additional limitations: Currently, only peptide ions identified using MS/MS can be quantified, and the doubling of peptide ion signals in isotopically labeled samples makes the spectra more complex (Mueller et al. 2008). These issues can be largely overcome by label-free LC-MS-based quantitative proteomic strategies, particularly if high mass accuracy MS is used (Rinner et al. 2007). Thereby, every detectable phosphopeptide ion signal can be tracked across multiple LC-MS feature maps, and each tracked peptide can be quantified (Mueller et al. 2007). As a result, the quantification process is applicable across a high number of samples.

Employing these advances, we established a novel, integrated experimental and computational pipeline for the label-free quantification of cellular phosphorylation patterns between a theoretically unlimited number of related samples. We applied this technique to compare the protein phosphorylation patterns in yeast cells upon cycloheximide or rapamycin treatment. This led to the identification of many phosphorylation sites that are presumably directly or indirectly targeted by TORC1. Repeating these assays using cells expressing TORC1-insensitive alleles of *TAP42* or *SCH9* we were able to accurately partition rapamycin-sensitive phosphoproteins to these two main TORC1 effector branches. These studies led to the observation that Maf1 is directly phosphorylated by Sch9, and that Sch9 regulates both Maf1 localization and binding to RNA Pol III. During the course of these experiments it became apparent that Sch9 also regulates RNA Pol I activity. Altogether this work reveals new effectors downstream from TORC1 and positions Sch9 as a central coordinator of protein synthesis.

Results

Sch9 and Tap42 act in parallel to mediate TORC1 signals

Two direct TORC1 effectors have been reported: Sch9 and Tap42. In order to test their functional relationship, we took advantage of the *SCH9*^{DE} and *tap42-11* alleles

that uncouple the encoded protein from upstream regulation by TORC1 (Di Como and Arndt 1996; Urban et al. 2007). In the TB50 yeast background *SCH9*^{DE} or *tap42-11* alone conferred only very slight resistance to rapamycin but, together, the two alleles showed strong synthetic rapamycin resistance with cells still growing, albeit slowly, in the presence of 200 nM rapamycin (~20 times the minimal lethal concentration for this strain; Fig. 1A). This argues that Sch9 and Tap42 function in parallel downstream from TORC1. The observation that the growth of *SCH9*^{DE} *tap42-11* cells is still slowed by the presence of rapamycin suggests the existence of additional direct TORC1 substrates like Sfp1 (Lempiainen et al. 2009). Beyond Tap42 and Sch9, the molecular events that couple TORC1 to its diverse range of downstream readouts remain largely uncharacterized.

Label-free quantitative phosphoproteomic screens

To tackle this challenge on a broad scale we employed a novel label-free quantitative phosphoproteomic approach as follows (Fig. 1B): Exponentially dividing cells were either treated with drug or mock-treated with drug vehicle. Subsequently, biochemical reactions were quenched. Proteins were extracted, enzymatically digested, and phosphopeptides were enriched using titanium dioxide-based affinity chromatography (Pinkse et al. 2004; Bodenmiller et al. 2007b). LC-MS feature maps, called phosphorylation patterns, were generated from each sample, and these phosphorylation patterns were aligned using the Superhirn algorithm (Mueller et al. 2007). Analysis of these alignments revealed phosphorylation events that were significantly regulated by a given treatment in a given genetic background.

Our specific strategy employed three different screens. First, we compared protein phosphorylation patterns of exponentially growing wild-type cells treated or mock-treated with cycloheximide, a translation elongation inhibitor that, by unknown mechanisms, hyperactivates TORC1 (Urban et al. 2007). In a second screen, protein phosphorylation patterns of wild-type cells and *SCH9*^{3E} cells (*SCH9*^{3E} is functionally identical to *SCH9*^{DE}) were compared both with and without rapamycin treatment. For the third screen, protein phosphorylation patterns of wild-type cells and *tap42-11* cells were similarly compared both with and without rapamycin treatment.

In total, 30 phosphorylation patterns were generated, containing 2256 distinct phosphopeptides mapping to 751 phosphoproteins (Supplemental Table S3; Supplemental File F1). The numbers of significantly up-regulated and down-regulated phosphopeptides in each screen are summarized in Supplemental Table S4. The summarized list of regulated phosphoproteins identified in both screens 2 and 3 and of proteins that were tested in independent phosphorylation assays is shown as Table 1. The comprehensive lists of regulated phosphopeptides are published as Supplemental Material (Supplemental File F2).

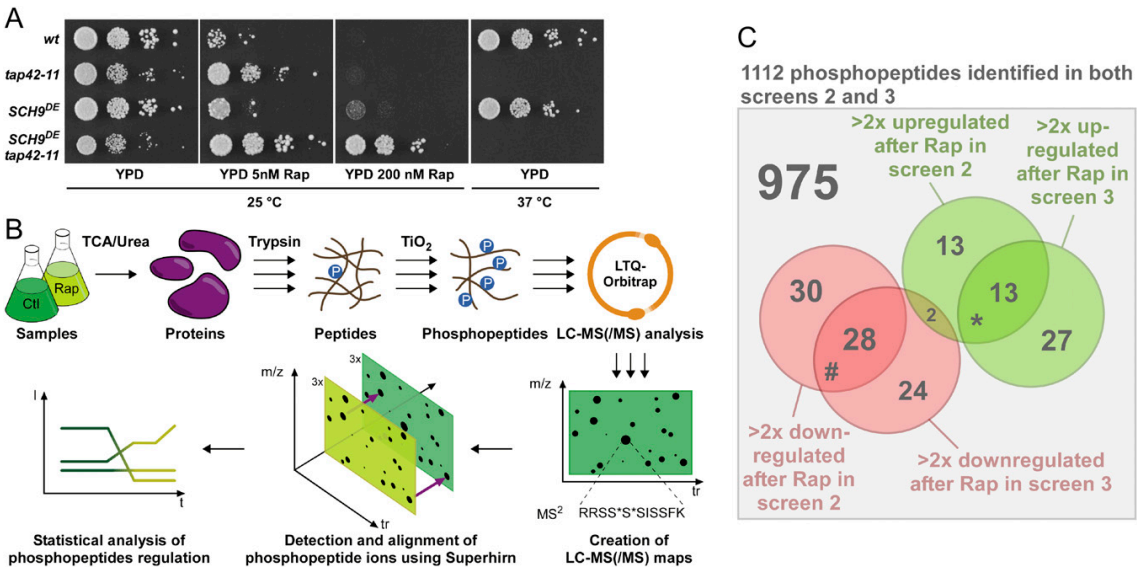


Figure 1. Label-free quantitative phosphoproteomic screens. (*A*) *TAP42* and *SCH9* act in parallel downstream from TORC1. Ten-fold serial dilutions of *sch9 tap42* cells complemented with indicated alleles of *SCH9* and *TAP42* and made prototroph with pAH149 were spotted onto the indicated media and incubated for 2–5 d at 25°C or 37°C. (Rap) Rapamycin. (*B*) Strategy for label-free quantitative phosphoproteomics. Triple arrows indicate steps performed in triplicate. (*C*) Venn diagram of phosphopeptides identified in both screens 2 and 3. Subsets of phosphopeptides found to be up-/down-regulated by rapamycin in each screen and their overlaps are shown. The overlap of phosphopeptides predicted to be down-regulated in screen 3 and up-regulated in screen 2 is not statistically significant (P = 0.25). P-values associated with the overlaps enrichment. (*) $P < 10^{-12}$; (#) $P < 10^{-24}$.

Table 1. *Summary of the phosphoproteomic screens*

ORF	Protein name	Regulation by rapamycin[a]	Regulation by cycloheximide[b]	Dependent on SCH9/TAP42	Reference[c]
Transcription, chromatin regulation					
YBL054W	Tod6	↘ (↘WB)	-	SCH9, TAP42	
YBL103C	Rtg3	↘	↘	SCH9, TAP42	Dilova and Powers 2006
YDR005C	Maf1	↘ (↘WB)	—	SCH9	Moir et al. 2006
YDR169C	Stb3	↘ (↘WB, ↘PS)	↗ (↗WB)	SCH9, TAP42	
YER040W	Gln3	↘	—		Beck and Hall 1999
YER088C	Dot6	↘ (↘WB)	—	SCH9	
YER169W	Rph1	↗ (↗WB)	— (−WB)		
YFL021W	Gat1	↘	—		Beck and Hall 1999
YIL038C	Not3	↘	↘	TAP42	
Translation					
YBR181C	Rps6b[d]	↘ ↗	—	SCH9	
YGR162W	Tif4631	↘	—		
YOR204W	Ded1	↗ (−WB, ↗PS)	↘		
YPL090C	Rps6a[d]	↘ ↗	—	SCH9	
YPR041W	Tif5	↗	—		
Signal transduction					
YDR028C	Reg1	↘ (−WB, −PS)	— (−WB)		
YHR082C	Ksp1	↘ (↘WB)	↗ (↗WB)	TAP42	
YHR205W	Sch9	↘	-	TAP42	Urban et al. 2007
YMR216C	Sky1	↘ (↘WB)	↗ (↗WB)		
YNL076W	Mks1	↘	↘	SCH9	Dilova et al. 2002
YPL180W	Tco89	↘	—	TAP42	
YPR185W	Atg13	↘	—	TAP42	Kamada et al. 2000
Transport					
YDR345C	Hxt3	↘	—	TAP42	
YJR001W	Avt1	↘ (−WB)	↗ (−WB)		
YNL321W	Vnx1	↗	—		
Metabolism					
YML035C	Amd1	↘	—		
YMR205C	Pfk2	↗	—		
YOL061W	Prs5	↗	—		
Miscellaneous/uncharacterized					
YBL051C	Pin4	— (↗WB)	↘ (↘WB)		
YCL011C	Gbp2[e]	↘ ↗	—	SCH9	
YCR077C	Pat1	↘	—		
YDL051W	Lhp1	↗ (−WB)	—	TAP42	
YDL173W	Par32	↘ ↗ (↗WB)	↗ (↘WB)	TAP42	
YDR348C		↘	—		
YIL047C	Syg1	↘ (−WB)	— (−WB)		
YIL135C	Vhs2	↘ ↗	↘	SCH9, TAP42	
YLR257W		↗	—		
YMR196W		↗	—		
YMR275C	Bul1	↗	—		
YNL004W	Hrb1[e]	↘ ↗	—	SCH9	
YNL265C	Ist1	↘ (−WB)	— (−WB)		
YOL060C	Mam3	↘	—		
YOR322C	Ldb19	↗	↘	TAP42	

List of phosphoproteins predicted to be regulated by rapamycin in both screens and/or tested to be rapamycin-sensitive by independent assays. Proteins are sorted according to their reported functions.

[a,b](↗) Up-regulated phosphopeptides; (↘) down-regulated phosphopeptides; (—) no regulated phosphopeptides. In parentheses, regulation observed in confirmation experiments [(WB) migration shift assay in Western blot; (PS) Phosphostaining; (↗) up-regulation of phosphorylation; (↘) down-regulation of phosphorylation; (—) no change in phosphorylation].

[c]Previous reports showing that the phosphorylation of the indicated proteins is regulated by rapamycin in a similar manner as in the phosphoproteomic screens.

[d]Rps6a and Rps6b have identical protein sequences.

[e]The peptides identified for Hrb1 and Gbp2 mapped to regions of the proteins that are identical in sequence and therefore could not be attributed to one or the other.

Quality assessment

To assess the quality of our data we first asked whether our screens were consistent with one another, and compared the phosphorylation patterns generated from independent experiments performed as part of our different screens. The overlap between the rapamycin-sensitive phosphopeptides of screens 2 and 3 was statistically significant (Fig. 1C).

Importantly, rapamycin-sensitive phosphorylation was found in proteins known to function downstream from TORC1, including Sch9, Gln3, Gat1, Atg13, Maf1, Mks1, and Rtg3 (Table 1). Thus, from our repeated independent experiments we observed consistent regulatory events, including those described in the literature. These results show the versatility of label-free quantitation, and confirm that our experimental and biological pipeline is highly reproducible.

In order to illustrate that our experimental and computational pipeline generates data of high quality, we chose to further validate the regulated phosphorylation sites (Table 1). We selected proteins from all three screens, with preference for those found regulated in two independent screens, showing opposite regulation by cycloheximide and rapamycin and/or displaying a high number of regulated phosphopeptides. HA-epitope-tagged versions of these selected proteins were expressed in yeast, and their migration in SDS-PAGE was monitored by Western blotting in the hope that the regulated phosphorylation sites would generate a migration "shift."

Dot6, Tod6, Ksp1, Sky1, and Stb3 were predicted to become hypophosphorylated after rapamycin treatment in the phosphoproteomic screens. Consistently, rapamycin treatment resulted in a faster migration of these proteins in SDS-PAGE (Fig. 2A,B), while cycloheximide had the inverse effect. Mutagenesis of some of the predicted phosphoserines in Dot6 and Tod6 to alanines resulted in variants whose migration was no longer altered by rapamycin treatment, indicating our protocol can accurately assign phosphorylation sites (Supplemental Fig. S1A). Ydl173w (renamed Par32 for phosphorylated after rapamycin, 32 kDa) and Rph1 were predicted to become hyperphosphorylated after rapamycin treatment (although some additional sites in Par32 were predicted to be hypophosphorylated after rapamycin treatment) and Pin4 was predicted to be dephosphorylated upon cycloheximide addition. Rapamycin treatment resulted in a slower migration of these proteins, while cycloheximide had the inverse effect. Syg1 and Avt1 were predicted to become hypophosphorylated after rapamycin treatment; however, no migration shift was observed after rapamycin or cycloheximide treatment for these two proteins.

As shifts in migration could result from other post-translational modifications we wished to confirm that the shifts that we had observed were indeed the result of changes in phosphorylation. To this end we treated selected immunoprecipitated phosphoproteins with λ-phosphatase in the absence or presence of phosphatase inhibitors. In each case, phosphatase treatment converted the protein to its fastest migrating species, confirming

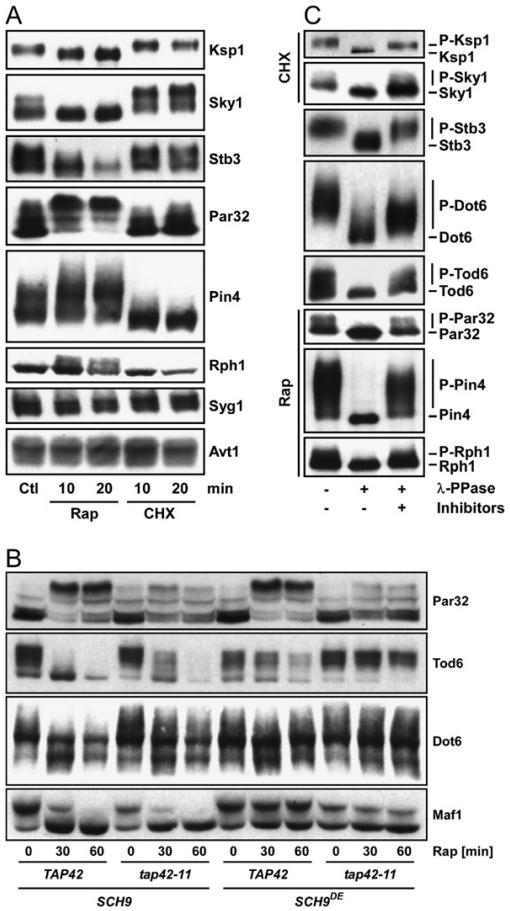


Figure 2. New TORC1 effectors. (*A*) Migration shift assays of proteins identified in the phosphoproteomic screens. Yeast cells expressing HA-tagged Ksp1, Sky1, Stb3, Par32, Pin4, Rph1, Syg1, and Avt1 were grown in YPD and treated with rapamycin or cycloheximide (CHX). Proteins were extracted under denaturing conditions and their SDS-PAGE migration was assayed by Western blotting. (*B*) Migration shift assays of phosphoproteins found to be regulated by *SCH9* or *TAP42*. Reporter plasmids expressing HA-tagged Par32, Tod6, Dot6, or Maf1 were transformed into cells of the indicated genotype. Cells were grown in YPD treated as indicated with rapamycin and assayed as in *A*. (*C*) Wild-type cells expressing HA-tagged Ksp1, Sky1, Stb3, Dot6, Tod6, Par32, Pin4, or Rph1 were grown in YPD and treated 15 min with rapamycin or cycloheximide where indicated. Proteins were extracted under native conditions and HA immunoprecipitates were incubated with λ phosphatase in the presence or absence of phosphatase inhibitors.

that differential phosphorylation was responsible for the observed migration shifts (Fig. 2C). In summary, eight of the 14 proteins we tested in migration shift assays (Fig. 2A,B; data not shown) gave the expected rapamycin-induced mobility shift as predicted in our phosphoproteomic screens. Immunoprecipitation experiments and staining for phosphorylated residues showed that Ded1,

which did not show any migration shift, becomes hyperphosphorylated after rapamycin treatment as predicted (Supplemental Fig. S1B). This suggests that at least some of the identified proteins that did not present observable migration shifts are nonetheless rapamycin-sensitive phosphoproteins.

The other objective of our phosphoproteome screens was to partition rapamycin-sensitive phosphorylation events amongst the two known TORC1 substrates, Tap42 and Sch9 (Supplemental Fig. S2; Supplemental Files F1, F2). Again, we used SDS-PAGE migration shift assays to determine if this objective had been met. For this analysis we chose Par32 (rapamycin-induced hyperphosphorylation predicted to be mediated by Tap42), Tod6 (rapamycin-induced hypophosphorylation predicted to be mediated by Sch9 and Tap42), Dot6 (rapamycin-induced hypophosphorylation predicted to be mediated by Sch9), and Maf1 (rapamycin-induced hypophosphorylation predicted to be mediated by Sch9). As predicted, the rapamycin-induced hyperphosphorylation of Par32 was largely blocked in *tap42-11* cells (Fig. 2B). Rapamycin-induced dephosphorylation of Tod6 and Dot6 was blocked in $SCH9^{DE}$ cells and was delayed in *tap42-11* cells (Tap42 dependence was predicted for Tod6 but not Dot6, which fell just under the applied cutoff for Tap42 regulation) (Fig. 2B). Lastly, Maf1 dephosphorylation showed the predicted Sch9 dependency (Fig. 2B).

Altogether, these control experiments demonstrate that our integrated experimental and computational pipeline allowed us to identify and quantify novel rapamycin-sensitive phosphorylation sites on a system-wide scale and to accurately partition these signaling events downstream from Tap42 and Sch9.

Novel TORC1 targets

We identified >100 novel TORC1-dependent phosphorylation events (Supplemental Table S4; Supplemental File F2), and we wanted to assess whether these phosphoproteins are important for TORC1 to regulate cell growth. Preliminary results suggest they are: 38 of 102 corresponding deletion strains tested gave a moderate to strong rapamycin phenotype (Supplemental Fig. S3; Supplemental Table S5). Thus, from our data we can anticipate novel functions for TORC1 and we can begin to explain how TORC1 signals to previously established readouts:

Previous studies demonstrated a role for TORC1 in starvation-induced developmental transitions (Cutler et al. 2001). TORC1 may mediate these transitions via Ksp1, a protein kinase that is required for filamentous growth in yeast (Bharucha et al. 2008).

Pin4 and Rph1 were shown to become hyperphosphorylated upon DNA damage (EM Kim et al. 2002; Pike et al. 2004). Curiously, we found that rapamycin induced a similar hyperphosphorylation suggesting cross-talk between TORC1 and DNA damage response pathways.

Our data suggest many new links to ribosome biogenesis. We found rapamycin-sensitive Sch9-dependent phosphorylation sites in Stb3 and Dot6/Tod6 (two homologous myb-like HTH transcription factors), and these proteins have been shown recently to function as transcriptional regulators of *ribi* genes (Liko et al. 2007; Badis et al. 2008; Zhu et al. 2009). Sky1 is a conserved Ser/Thr kinase that phosphorylates pre-mRNA splicing factors of the SR family (Siebel et al. 1999). Given the prevalence of introns in *RP* genes, this could suggest that TORC1 also plays a more direct role in ribosome biogenesis.

We chose to focus our attention on yet another protein implicated in ribosome biogenesis regulation. Specifically, we were intrigued that Maf1, a conserved repressor of RNA Pol III, was predicted to be regulated by TORC1 via Sch9 (Fig. 2B), as this regulation was thought to occur via PP2A phosphatases (Oficjalska-Pham et al. 2006).

Maf1 is a direct target of Sch9

Previously, we demonstrated that Sch9 phosphorylated two serines near the C terminus of Rps6 (Urban et al. 2007). Notably, the sequence surrounding the phosphoserines now identified in Maf1 bears a striking similarity to this S6 sequence (Fig. 3A). Maf1 was reported to be dephosphorylated after rapamycin treatment, but the mechanism by which TORC1 regulates Maf1 phosphorylation was not determined (Oficjalska-Pham et al. 2006; Roberts et al. 2006). Although PP2A family protein phosphatases have been implicated in Maf1 dephosphorylation, we confirmed earlier reports (Willis et al. 2004; Willis and Moir 2007) that Tap42 does not play a role in rapamycin-induced Maf1 dephosphorylation (Fig. 2B).

We wished to determine if Sch9 inhibition alone could cause dephosphorylation of Maf1. In our hands, *SCH9* deletion mutants grow very slowly but rapidly accumulate suppressive mutations potentially confounding conclusions derived from these strains. We therefore took advantage of a previously described analog-sensitive allele of *SCH9* ($sch9^{as}$) encoding a protein that can be specifically inhibited by the bulky ATP analog 1NM-PP1 (Jorgensen et al. 2004). Addition of 1NM-PP1 to $sch9^{as}$, but not wild-type cells, resulted in a rapid dephosphorylation of Maf1 (Fig. 3B).

As Maf1 was reported to be regulated by the RAS–PKA pathway and to be phosphorylated in vitro by PKA (Moir et al. 2006), we wished to further explore the relative contributions of Sch9 and PKA to Maf1 phosphorylation. PKA is encoded by three genes in yeast (*TPK1*, *TPK2*, and *TPK3*) and is regulated by glucose in parallel to the TOR pathway (Dechant and Peter 2008). Deletion of the three *TPK* genes is lethal but can be rescued by the deletion of *YAK1* (Garrett and Broach 1989). We found that Maf1 is still phosphorylated in *tpk1 tpk2 tpk3 yak1* cells, and this phosphorylation is still sensitive to rapamycin treatment (Fig. 3C). In contrast to the results obtained with $sch9^{as}$ cells, Maf1 phosphorylation was only slightly affected by addition of 1NM-PP1 to $tpk1^{as}$ $tpk2^{as}$ $tpk3^{as}$ cells (Supplemental Fig. S4A,B), suggesting a minor role of PKA under these conditions. Maf1 dephosphorylation during transit through diauxic shift did not show any significant differences in *tpk1 tpk2 tpk3 yak1* versus wild-type cells (Supplemental Fig. S4C). In addition, Maf1 dephosphorylation following nitrogen starvation (Fig. 3D) was partially

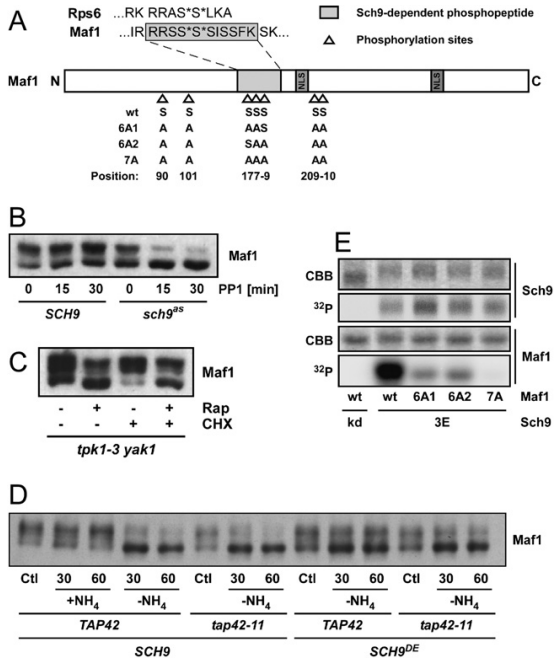


Figure 3. TORC1 regulates Maf1 phosphorylation via SCH9, independently of PKA. (*A*) Maf1 schematic. Maf1 features including phosphorylation sites and NLSs are pictured. Serines predicted to be phosphorylated in the MS data are followed by asterisks as are the Sch9 target residues in the C terminus of Rps6. The various alanine-substituted versions of Maf1 used in the kinase assays shown in *D* are summarized *below* the scheme. (*B*) Sch9 inhibition leads to Maf1 dephosphorylation. (*C*) TORC1 regulates Maf1 phosphorylation independently of PKA. (*B*,*C*) Protein extracts were prepared from cells of the indicated genotype following treatment (15 min in *C*) with the indicated drugs (PP1: 1NM-PP1). Phosphorylation of Maf1-3HA was determined by SDS-PAGE and Western blotting. (*D*) Sch9 couples nitrogen-dependent signals to Maf1. Prototroph cells of the indicated genotype were grown to exponential phase in SD, filtered, and resuspended in control ($+NH_4$) or in ammonium-deprived medium ($-NH_4$). Samples were taken at the indicated time points and analyzed for Maf1 phosphorylation by Western blotting. (*E*) Sch9 phosphorylates 7 serines in Maf1 in vitro. Maf1 mutants, purified from *Escherichia coli*, were tested as substrates for GST-Sch9^{3E} purified from yeast. GST-Sch9^{kd} is a point mutant lacking catalytic activity and was used as a negative control. Reactions were resolved by SDS-PAGE, proteins were stained with Coomassie (CBB) and the dried gel was analyzed for ^{32}P incorporation.

blocked in cells expressing the *SCH9^{DE}* allele. Altogether, these data suggest that Sch9 is the major kinase upstream of Maf1 in exponentially growing cells.

While this manuscript was in preparation, Lee et al. (2009) reported that Sch9 phosphorylates Maf1; but, in contrast to our results, their work suggested that Sch9 and PKA perform equally important roles in Maf1 regulation. We believe that this discrepancy is due to protocol differences in the handling of yeast cells prior to protein extract. In our studies, whenever possible, TCA was added to growing cultures to quench all enzymatic activity prior to further manipulations. In contrast, Lee et al. (2009) cooled the cells on ice prior to lysis. We found that this cooling step elicits a PKA-dependent rephosphorylation of Maf1 (Supplemental Fig. S8B). This observation suggests a more prominent role of PKA in Maf1 phosphorylation at lower temperatures and explains the overestimated importance in Maf1 regulation at 30°C assigned to PKA by Lee et al. (2009).

Next, we asked if Sch9 could directly phosphorylate Maf1 in vitro. We found that Sch9^{3E}, but not a kinase-dead mutant, was able to phosphorylate purified recombinant Maf1 (Fig. 3E). Maf1 contains six motifs fitting the R[R/K]xS consensus, which is often attributed to AGC family kinases (S90, S101, S177, S178, S209, and S210). The phosphopeptide identified in our phosphoproteomic screens (Fig. 3A) contains two overlapping copies of this motif. Unfortunately, neither the Sequest algorithm used to annotate the tandem mass spectra nor manual inspection allowed us to determine with certainty which of the four consecutive serines in this peptide were phosphorylated. To resolve this issue we generated various alanine substituted versions of Maf1 (Fig. 3A) and assayed the ability of Sch9 to phosphorylate these proteins in vitro. First we replaced all six serines fitting the R[R/K]xS consensus with alanine. This version (Maf1^{6A1}) was still phosphorylated in vitro, albeit very poorly, by Sch9 (Fig. 3E). Based on the phosphopeptide that we identified in our screen we chose to substitute an additional serine residue (S179) generating a version of Maf1 we refer to as Maf1^{7A}. Maf1^{7A} was not a substrate for Sch9 in vitro, suggesting that we mapped all of the Maf1 residues phosphorylated by Sch9.

TORC1 regulates RNA Pol III through Sch9

Having established that TORC1, via Sch9, regulates Maf1 phosphorylation we next asked whether this cascade is physiologically important for the regulation of RNA Pol III function. To begin, we confirmed that rapamycin treatment causes a dramatic reduction in tRNA and 5S rRNA synthesis as determined by ^3H-uracil pulse labeling (Fig. 4A). This drop in RNA Pol III activity was largely blocked in cells expressing *SCH9^{DE}*, whereas *tap42-11* alone seemed to play little if any role in this process (Fig. 4A). These observations demonstrate that Sch9 influences RNA Pol III activity.

To extend these observations we wished to determine if Sch9 signals to RNA Pol III via Maf1. 1NM-PP1 addition to *sch9^{as}* cells resulted in a rapid inhibition in tRNA and 5S rRNA synthesis (Fig. 4B). *MAF1* deletion abrogated the 1NM-PP1-induced reduction of tRNA synthesis but seemingly did not abrogate the reduction of 5S rRNA synthesis. An explanation for this result is explored below. Pre-tRNAs are rapidly processed in exponentially growing cells, and their abundance can thus be used to infer RNA Pol III activity. Quantitative RT–PCR analyses of the pre-tRNA^{Pro} levels were consistent with the ^3H-uracil pulse labeling experiments and confirmed the epistasis between *SCH9* and *MAF1* (Supplemental Fig. S5).

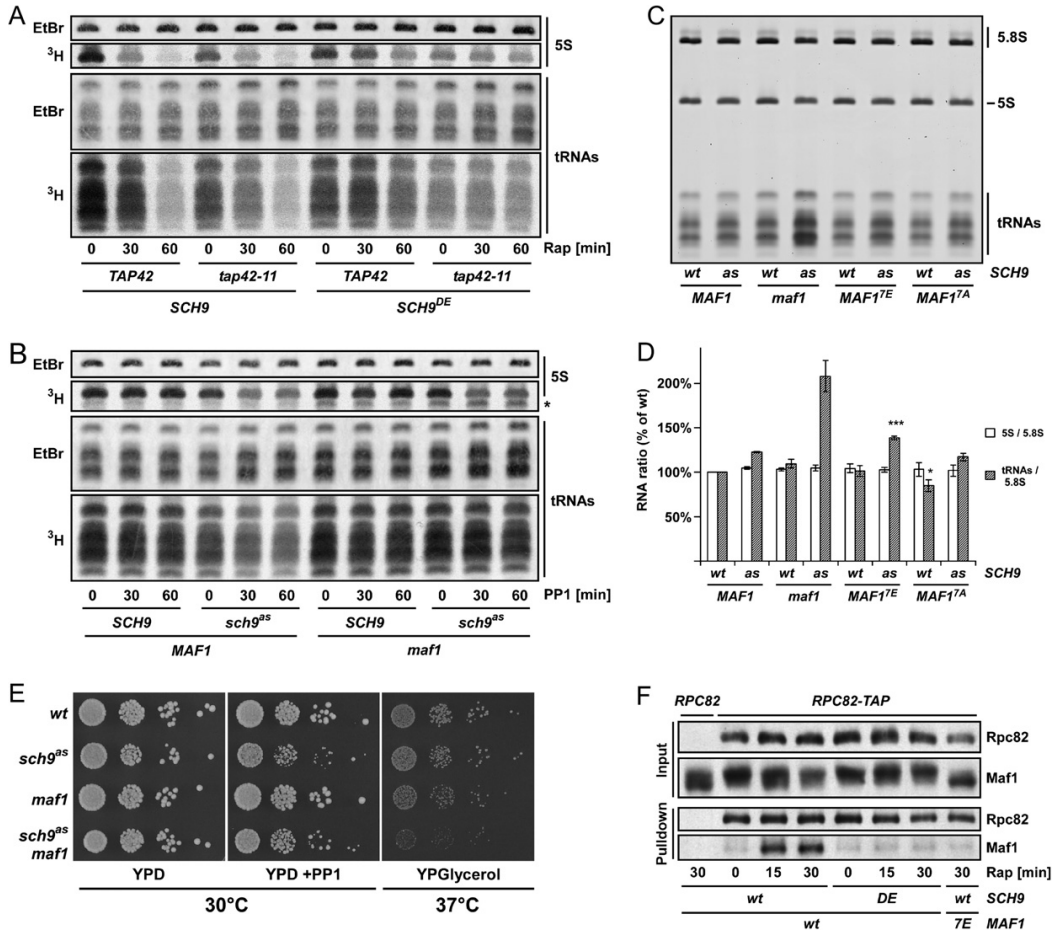


Figure 4. TORC1 regulates RNA Pol III via Sch9 and Maf1. (*A*) Rapamycin inhibits 5S rRNA and tRNA synthesis via *SCH9*. (*B*) Sch9 inhibition leads to a Maf1-dependent inhibition of tRNA synthesis. (*A*,*B*) RNA synthesis in cells of the indicated genotype following the indicated drug treatment was determined by metabolic labeling with ^{3}H-uracil. Total RNA loaded was determined by staining with ethidium bromide (EtBr). Asterisk (*) indicates an unstable RNA species that accumulates in *maf1* cells. (*C*) *MAF1* phosphorylation regulates tRNA levels. Cells of the indicated genotype were grown in SC −URA −LEU −TRP −HIS 0.2% Gln 300 nM 1NM-PP1 to log phase (OD_{600} < 0.8) and total RNA was extracted. Total levels of the 5S and 5.8S rRNA and tRNA were assayed by PAGE and EtBr staining. (*D*) Quantification of *C* and two other independent experiments. 5S:5.8S and tRNA:5.8S ratios were calculated and plotted relative to untreated wild-type control. (*) $P < 0.05$; (***) $P < 0.001$ versus wild-type control. (*E*) Genetic interactions between *SCH9* and *MAF1*. Ten-fold dilutions of the indicated strains were spotted and grown on YPD ± 1NM-PP1 (2 d, 30°C) or on YPGlycerol (4 d, 37°C). (*F*) Sch9 regulates Maf1 association with RNA Pol III. Interaction of Maf1-3HA with RNA Pol III was assessed by Rpc82-TAP pull-downs followed by SDS-PAGE and Western blotting. Relevant genotypes and rapamycin treatments are indicated.

These observations prompted us to determine whether the deregulation of tRNA synthesis observed in our *sch9^{as} maf1* mutant could lead to an altered abundance of tRNA relative to other RNA species. To this end we compared 5.8S, 5S, and tRNA levels in exponentially growing cells in the presence or absence of Sch9 and/or Maf1 function. Although total 5S rRNA levels relative to RNA Pol I-derived 5.8S rRNA were unchanged in any of the strains examined, the tRNA:5.8S rRNA ratio was approximately twofold higher in 1NM-PP1-treated *sch9^{as} maf1* cells compared with wild-type cells (Fig. 4C,D). An explanation for this result is explored below. To evaluate the importance of Maf1 phosphorylation in RNA Pol III regulation we employed Maf1 mutants where all seven Sch9 phosphorylated residues were replaced with either glutamate (Maf1^{7E}) or alanine residues (Maf1^{7A}). Inhibition of Sch9 in cells expressing Maf1^{7E} leads to tRNA accumulation intermediate to that observed in *MAF1* and *maf1* cells; cells expressing the Maf1^{7A} mutant had reduced basal levels of tRNA (Fig. 4C,D). ^{3}H-uracil pulse labeling experiments following Sch9 inhibition and rapamycin treatment in *MAF1*, *maf1*, *MAF1^{7A}*, and *MAF1^{7E}*

strains similarly demonstrate that Maf1 phosphosite mutants are partially functional (Supplemental Fig. S6A,B). These results are consistent with Sch9 regulating RNA Pol III via Maf1, but suggest that Maf1 phosphorylation plays only a partial role in this regulation, implying that Sch9 targets additional factors to regulate RNA Pol III albeit in a Maf1-dependent manner.

Additional genetic observations further support our observations. In the presence of 1NM-PP1, *sch9^{as}* cells grew slowly, dividing every 150 min ± 8 min (compared with 103 min ± 4 min for wild-type cells). This slow growth rate was slightly, but significantly, improved to 135 min ± 5 min by deletion of *MAF1*, suggesting that reduced RNA pol III activity is one of multiple growth-limiting consequences resulting from loss of Sch9 function. Conversely, we observed a synthetic growth defect when *sch9^{as} maf1* cells were grown at 37°C on the nonfermentable carbon source glycerol (Fig. 4E). This result fits with the previously proposed model that accumulation of tRNA in *maf1* cells is detrimental to mitochondrial function (Boguta et al. 1997).

Sch9 regulates Maf1 localization and association with the RNA Pol III machinery

How does phosphorylation by Sch9 alter the ability of Maf1 to inhibit RNA Pol III activity? We explored two potential mechanisms by which Sch9 might regulate Maf1: Maf1 localization and the capacity of Maf1 to bind to RNA Pol III. Maf1 was both nuclear and cytoplasmic in our strain background and promptly accumulated in the nucleus upon rapamycin treatment (Supplemental Fig. S7A,B). Sch9^{DE} did not block the rapamycin-induced nuclear accumulation; but, probably due to the hypomorphic nature of these alleles (Jorgensen et al. 2004; Urban et al. 2007), *SCH9^{DE}* and *sch9^{as}* cells showed increased basal nuclear accumulation of Maf1 that could be further enhanced in *sch9^{as}* cells by 1NM-PP1 treatment (Supplemental Fig. S7A–D). From these results, we propose that Maf1 phosphorylation by Sch9 contributes but is not sufficient to prevent Maf1 nuclear localization. This hypothesis is consistent with previous models (Moir et al. 2006) suggesting that TORC1 regulates the two redundant nuclear localization sequences (NLSs) found in Maf1 (Fig. 3A) independently; activity of the N-terminal NLS was proposed to be regulated via phosphorylation (i.e., Sch9-dependent) while the C-terminal NLS is regulated independently of phosphorylation (i.e., Sch9-independent).

In contrast to localization, the ability of Maf1 to directly bind to RNA Pol III appears to be a more important mechanism by which Maf1 phosphorylation regulates RNA Pol III activity (Oficjalska-Pham et al. 2006; Roberts et al. 2006). Therefore we wished to test if phosphorylation by Sch9 alters the ability of Maf1 to interact with the RNA Pol III subunit Rpc82. Using a coprecipitation assay we confirmed that rapamycin treatment strongly increases the association of Maf1 with Rpc82. Importantly, we found that Sch9^{DE} blocked this rapamycin-induced interaction (Fig. 4F). Consistent with this observation, the Maf1^{7E} mutant did not associate with Rpc82 upon rapamycin treatment (Fig. 4F), while the Maf1^{7A} mutant showed constitutive interaction in untreated wild-type cells (Supplemental Fig. S8A). We were not, however, able to detect an induction of the interaction upon Sch9 inhibition in the TB50 genetic background (Supplemental Fig. S8A), which is likely due to Maf1 rephosphorylation that occurs during cooling prior to nondenaturing protein extraction (Supplemental Fig. S8B). We observed a small induction of the interaction upon Sch9 inhibition in the W303 background that could be further enhanced by concomitant inhibition of PKA (Supplemental Fig. S8C). Taken together, these results demonstrate that Sch9 (and PKA in some conditions) regulates the capacity of Maf1 to bind RNA Pol III.

Sch9 mediates the TORC1 signal to RNA Pol I

As noted above (Fig. 4C,D), we observed that the ratio of tRNA:5.8S rRNA is increased in 1NM-PP1-treated *sch9^{as} maf1* cells relative to untreated *sch9^{as} maf1* cells. It was not clear why tRNA should be accumulated relative to 5.8S rRNA in these exponentially growing cells; i.e., when rRNA expression should be fully derepressed. We rationalized that this result could be explained if both RNA Pol I and III activities are reduced upon Sch9 inhibition, with *MAF1* deletion suppressing only the latter. Thus, tRNA synthesis was not up-regulated per se, but rather RNA Pol I-dependent rRNA synthesis was decreased in these cells. This hypothesis could also explain the apparent failure of *MAF1* deletion to rescue the decrease in RNA Pol III-dependent 5S rRNA synthesis resulting from Sch9 inhibition (Fig. 4B): 5S rRNA transcription, we speculate, is rescued by *MAF1* deletion. However, because RNA Pol I activity is reduced, 5S rRNA is produced in excess relative to other rRNAs and, consequently, is unstable and rapidly degraded. The shorter unstable RNA species whose levels increased when Sch9 activity was inhibited in *maf1* cells (asterisk in Fig. 4B) fits well with this notion. These deductions prompted us to test if TORC1 regulates RNA Pol I via Sch9.

We first confirmed that rapamycin treatment results in a rapid decline in the synthesis of 25S, 18S, and 5.8S rRNAs, as judged by ^{3}H-uracil pulse labeling assays (Fig. 5A). This effect was dramatically blocked in *SCH9^{DE}* cells but unaltered in *tap42-11* cells. Interestingly, ^{3}H-uracil incorporation into these rRNAs was virtually insensitive to rapamycin in *SCH9^{DE} tap42-11* cells. Consistently, addition of 1NM-PP1 to *sch9^{as}* cells resulted in a rapid, Maf1-independent, decrease in the synthesis of 25S, 18S and 5.8S rRNAs (Fig. 5B).

These results clearly demonstrated that Sch9 is indeed important in regulating the synthesis of RNA pol I-derived rRNA species and beg the question: How? Sch9 could regulate RNA pol I transcription or rRNA processing or a combination thereof. To begin to discriminate between these possibilities we first determined the relative amounts of unstable rRNA species to try to gauge the flux from pre-rRNAs to mature end products. Rapamycin treatment and Sch9 inhibition led to decreased rRNA processing as judged by dramatically increased 27S:25S

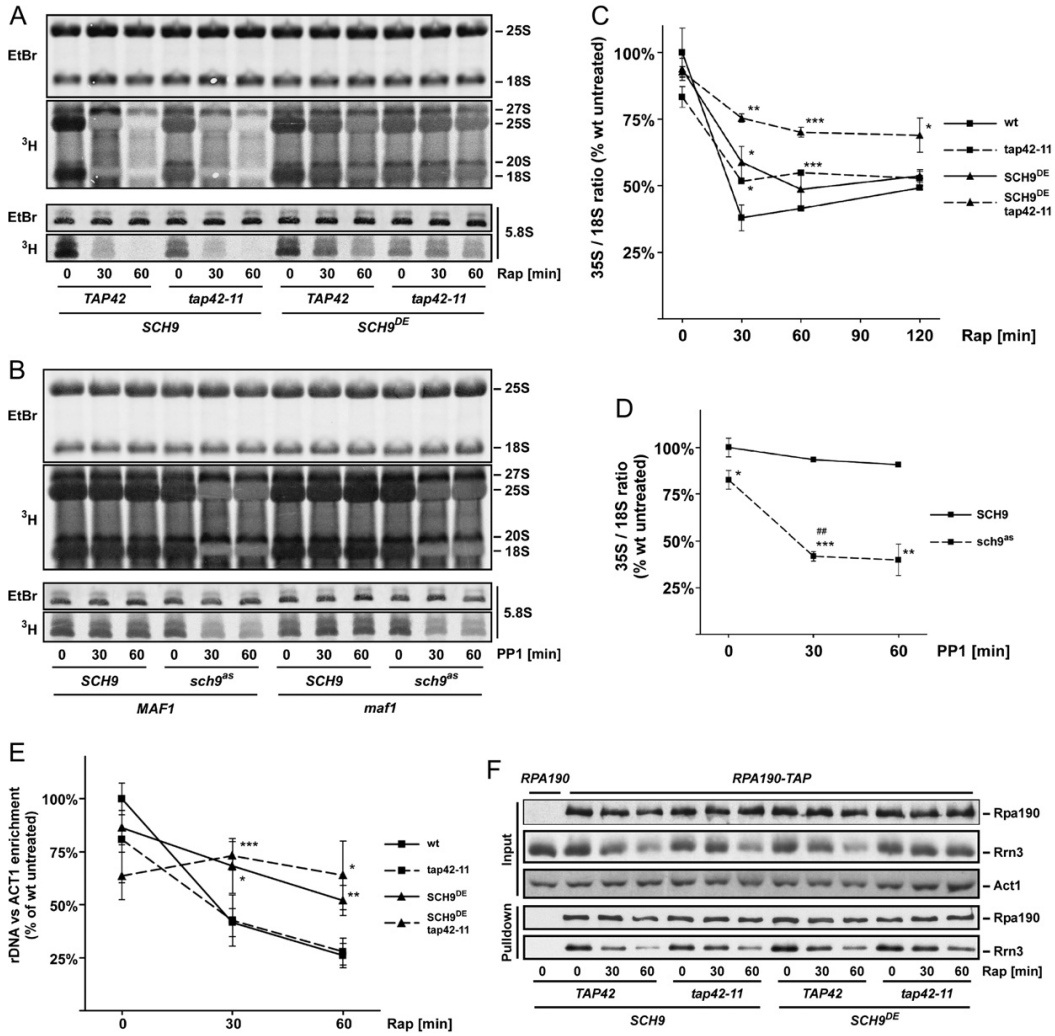


Figure 5. TORC1 regulates RNA Pol I via Sch9. (*A*) Rapamycin treatment decreases the processing/expression of RNA pol I-derived rRNA species. (*B*) Sch9 inhibition decreases the processing/expression of RNA pol I-derived rRNA species. (*A*,*B*) Synthesis/processing of rRNA was assayed by metabolic labeling with ^{3}H-uracil. Total RNA loaded was determined by staining with ethidium bromide (EtBr). (*C*) Rapamycin treatment decreases 35S pre-rRNA synthesis. (*D*) Sch9 inhibition decreases 35S pre-rRNA synthesis. (*C*,*D*) 35S and 18S rRNA levels were determined by primer extension—gels are shown in Figure S9—and their ratios were plotted. Values are means of three independent experiments ± SD. (*E*) RNA Pol I recruitment at the rDNA locus depends on Sch9. Association of RPA190-13myc with the rDNA locus was determined by ChIP. Values are means of four independent experiments ± SD. Statistical confidences for *C*–*E*: (*) $P < 0.05$; (**) $P < 0.01$; (***) $P < 0.001$ versus wild-type control; (##) $P < 0.01$ versus untreated isogenic control. (*F*) Sch9 does not regulate Rrn3–RNA Pol I interaction. Association of Rrn3-5HA with Rpa190-TAP was assayed using TAP pull-downs and SDS-PAGE/Western blotting. (*A*–*F*) Relevant genotypes and rapamycin/1NM-PP1 treatment times are indicated.

and 20S:18S ratios (Fig. 5A,B). Rapamycin-induced processing defects were largely blocked by the *SCH9^{DE}* allele. Thus, Sch9 activity promotes rRNA processing.

We also explored if Sch9 might regulate RNA Pol I activity by several approaches. First, we used primer extension assays to determine the relative abundance of the short-lived 35S pre-rRNA. In wild-type cells, 35S levels dropped approximately threefold when assayed after 30 min of rapamycin treatment (Fig. 5C). *SCH9^{DE}* and *tap42-11* cells showed slight, but significant resistance, while *SCH9^{DE} tap42-11* cells showed pronounced resistance to this rapamycin-induced drop in 35S pre-rRNA levels. Consistent with these observations, addition of 1NM-PP1 to *sch9^{as}* cells resulted in an approximately twofold drop in 35S pre-rRNA levels (Fig. 5D). Although it is difficult to separate processing effects from RNA pol I

activity, these observations suggest that Sch9 and Tap42 each play a role in regulating RNA Pol I transcription.

Next, we used chromatin immunoprecipitations (ChIPs) to more directly evaluate a role for Sch9 and/or Tap42 in RNA Pol I activity. ChIPs with Rpa190, the catalytic subunit of RNA Pol I, showed that RNA Pol I occupancy at the *rDNA* locus decreased more than twofold following 30 min of rapamycin treatment (Fig. 5E). Mirroring the ^{3}H-uracil labeling assays, this rapamycin-induced eviction of RNA Pol I from the *rDNA* locus was strongly blocked in $SCH9^{DE}$ cells and completely blocked in $SCH9^{DE}$ *tap42-11* cells. Thus, TORC1 promotes the recruitment of RNA Pol I to the *rDNA* locus primarily via Sch9.

Previously, Claypool et al. (2004) had proposed that TORC1 influenced RNA Pol I recruitment to the *rDNA* locus by promoting an interaction between RNA Pol I and Rrn3, an essential initiation factor. We therefore asked if Sch9 and/or Tap42 impinge on this Rrn3–RNA Pol I interaction by pull-down of TAP-tagged Rpa190. This appears not to be the case: The rapamycin-induced dissociation of Rrn3 and RNA Pol I in $SCH9^{DE}$ *tap42-11* cells was essentially indistinguishable from the dissociation observed in wild-type cells (Fig. 5F). These data suggest that the dissociation of Rrn3 from RNA Pol I is not the primary mechanism by which TORC1 inhibition causes a reduction in *rDNA* transcription. We do note, however, that Rrn3 levels drop during rapamycin treatment, and that this is blocked in $SCH9^{DE}$ *tap42-11* cells.

Discussion

Sch9 is a central coordinator of protein synthesis

Previous work (Upadhya et al. 2002) had shown that phosphorylation of the RNA Pol III inhibitor Maf1 is regulated downstream from TORC1 and, consistently, a Maf1 phosphopeptide was found to be down-regulated after rapamycin treatment in our phosphoproteomic screens. However, our screens further predicted that Maf1 phosphorylation is regulated by TORC1 via Sch9. Subsequent biochemical studies demonstrated that Maf1 is likely a direct substrate of Sch9 and genetic experiments demonstrated that RNA Pol III down-regulation upon Sch9 inhibition is Maf1-dependent. At the molecular level, we found that Maf1 phosphorylation by Sch9 was important for Maf1 association with RNA Pol III. However, RNA Pol III activity is still sensitive to rapamycin in cells expressing a "phosphomimetic" version of Maf1 ($Maf1^{7E}$). Thus, we believe that Sch9 must target a factor(s) in addition to Maf1 to regulate RNA Pol III activity.

Expectedly, *MAF1* deletion, which mostly affects tRNA levels, did not suppress the *sch9* growth phenotype. However, it would be interesting in the future to study the impact of Maf1 on other Sch9 phenotypes such as cell size regulation and longevity (Jorgensen et al. 2004; Kaeberlein et al. 2005). In particular, regulation of tRNA and especially initiator $tRNA^{Met}$ levels by Maf1 could affect translation and, via Gcn4, longevity as recently observed (Steffen et al. 2008).

During the course of these studies we found that, in addition to RNA Pol III, Sch9 also regulates the synthesis of RNA Pol I transcripts. Specifically, Sch9 promotes both processing of rRNA species and recruitment of RNA Pol I to the *rDNA* locus. rRNA processing could be an indirect function of Sch9 as Sch9 controls the expression of many rRNA processing factors in the *ribi* regulon (Jorgensen et al. 2004). We do not know the mechanism by which Sch9 promotes recruitment of RNA Pol I to the *rDNA* locus but it does not appear to involve the association of RNA Pol I with its initiation factor Rrn3. Interestingly, RNA Pol I activity was proposed to be determinant for the expression of other ribosomal components (RPs and 5s rRNA), which could suggest that Sch9 regulates *RP* genes expression indirectly via RNA Pol I activity (Laferte et al. 2006).

Sch9 was thought previously to be the ortholog of mammalian Akt. However, we proposed recently that Sch9 functions more similarly to mammalian ribosomal S6 kinase (S6K1) (Urban et al. 2007). This present study strengthens the functional similarities between Sch9 and S6K1: Both Sch9 and S6K1 have now been shown to regulate the activities of the three nuclear RNA polymerases (Zhang et al. 2005; Woiwode et al. 2008). Recent studies have shown that elevated RNA Pol III transcription is necessary, or in some cases sufficient, for cellular transformation (Johnson et al. 2008; Marshall et al. 2008). Extrapolating from our results, it will be of interest to determine if dysregulation of S6K1 and/or Maf1 likewise contribute to cellular transformation.

In summary, the present study, together with previous work, demonstrates that Sch9 regulates the activities of all three nuclear RNA polymerases. In addition, we showed previously that Sch9 also regulates translation initiation (Urban et al. 2007). Thus, Sch9 appears to play a central role in coupling environmental cues to the coordinated expression, assembly, and activity of the protein synthesis machinery (Fig. 6). In addition to Maf1, our characterization of the rapamycin-sensitive phosphoproteome also uncovered other TORC1/Sch9 targets implicated in ribosome biogenesis; namely, Stb3 and Dot6/Tod6, which have been shown previously to respectively bind RRPE and PAC elements in *ribi* gene promoters (Liko et al. 2007; Badis et al. 2008; Zhu et al. 2009). We found Sch9-dependent rapamycin-sensitive phosphorylation sites in Stb3 and Dot6/Tod6, many of which fit the R[R/K]xS consensus motif, suggesting that these proteins could be directly phosphorylated by Sch9. Western blot analyses confirmed that Dot6 and Tod6 are indeed phosphorylated downstream from Sch9 (Fig. 2B), and preliminary experiments suggest that Sch9 signals antagonize the ability of Stb3, Dot6, and Tod6 to inhibit *ribi* gene expression (Fig. 6; Supplemental Figs. S1A, S3; data not shown).

Label-free quantitative phosphoproteomic screens

The integrated experimental and computational framework that we present in this work enables relative quantification of phosphorylation patterns. The procedure

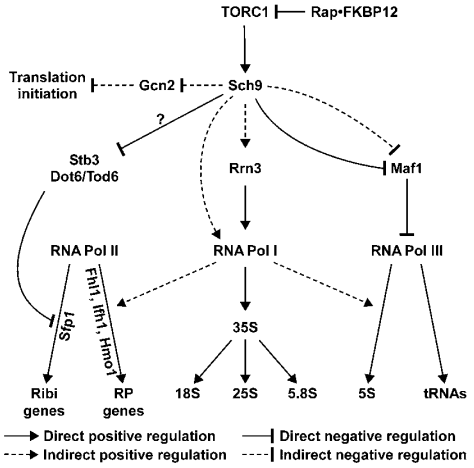


Figure 6. Model of TORC1 signaling highlighting the central role that Sch9 plays in coordinating the expression, assembly and activity of the protein synthesis machinery. See the Discussion for details.

is technically robust and sensitive, and the data acquired is accurate and, as we illustrate in this manuscript, highly reproducible. The main advantage of our method is that it allows us to quantify thousands of phosphorylation sites between, in principle, an unlimited number of samples or biological states. In addition, no a priori knowledge of phosphopeptide ions is required for quantitation as targeted LC-MS/MS methods can be employed to identify regulated ions of interest (Schmidt et al. 2008). This improves the sensitivity and especially the achievable throughput compared with quantification based on isotope labeling.

Although our data identified many of the rapamycin-sensitive phosphorylation events described in the literature, many were missed. Indeed, high coverage of a given phosphoproteome remains a major challenge in current phosphoproteomics (Bodenmiller et al. 2007b). Nevertheless, this study demonstrates that even with an incomplete coverage of a phosphoproteome our experimental and computational pipeline can elucidate novel and important biological phenomena as demonstrated with the characterization of Maf1 phosphorylation.

Materials and methods

Yeast cultures and assays

Saccharomyces cerevisiae strains and plasmids are described in Supplemental Tables S1 and S2, respectively. Strains were constructed according to standard protocols. *SCH9* and *TAP42* were deleted in diploid strains and complemented with plasmids encoding wild-type alleles before sporulation. Wild-type alleles were subsequently replaced with mutant alleles in haploids by plasmid shuffling.

Unless specified otherwise, rapamycin was used at 200 nM (from a 1 mM stock solution in 90% ethanol, 10% Tween-20), 1NM-PP1 at 200 nM (from a 1 mM stock in DMSO), and cycloheximide at 25 μg/mL (from a 10 mg/mL stock solution in H_2O).

Label-free phosphoproteomics

Cells were grown in SC −LEU 0.2% Gln at 23°C to OD_{600} 0.8 and subjected to the indicated treatments for 30 min. All biochemical activities were then quenched by the addition of trichloroacetic acid and proteins were extracted under denaturing conditions. Three 400-μL aliquots for each condition were processed separately for disulfide bond reduction, cysteine alkylation, trypsin digestion, and phosphopeptide enrichment as described in more detail in the Supplemental Material.

The phosphopeptides were separated by reverse phase chromatography on an Eksigent nano-LC system and were analyzed on a hybrid LTQ-Orbitrap mass spectrometer (Thermo Electron Corporation) interfaced with a nano-electrospray ion source as detailed in the Supplemental Material. The LC-MS/MS data was searched against the SGD yeast protein database as described in the Supplemental Material.

For the detection of the regulated features (peak picking and integration of the area from the LC-MS data, alignment of features over multiple runs) the SuperHirn version 2.0 algorithm (Mueller et al. 2007; Rinner et al. 2007) was used. Of note, as we used a label-free approach, the peptide sequence information from all LC-MS/MS runs was usable to assign the LC-MS features present in the SuperHirn output file, called MasterMap (the relevant parameters used are published as Supplemental Material). The MasterMap was post-processed as follows: Of all of the phosphopeptide features the deconvoluted masses were computed and the peak areas of phosphopeptides present in different charge states were merged. Based on these areas, the statistical significance was computed using a *t*-test as implemented in the Corra software environment (Brusniak et al. 2008).

Phosphopeptides were considered to be cycloheximide- or rapamycin-sensitive if their abundance relative to untreated cells was altered twofold or more (*P*-value, <0.05) by cycloheximide or rapamycin treatment, respectively. Rapamycin-sensitive phosphopeptides were considered to be Sch9-dependent after fulfilling two selection criteria. First, rapamycin-induced changes in wild-type cells had to be blunted twofold or more in $SCH9^{3E}$ cells. Second, the abundance of a phosphopeptide reduced (or induced) by rapamycin treatment in wild-type cells had to be ≥1.5-fold higher (or lower) in both untreated and rapamycin-treated $SCH9^{3E}$ cells compared with rapamycin-treated wild-type cells. The same criteria were used to evaluate Tap42 dependence. All regulated phosphopeptide ions corresponding to the phosphorylation sites of interest were validated.

Data availability

All MS^2 information will be made available via the Phosphopep database (http://www.phosphopep.org) (Bodenmiller et al. 2007a). The raw data in the mzXML format can be downloaded from the Peptide Atlas Web page at http://www.peptideatlas.org/repository (Desiere et al. 2006).

Copurification assays

Precultures grown in plasmid-selective synthetic medium were diluted in YPD and grown to OD_{600} 0.7–1.0. One-hundred-milliliter aliquots were treated as described in the text and processed for native protein extraction as described in the Supplemental Material. Protein concentrations were normalized and aliquots were removed to control for input. TAP pull-downs were performed for 2 h at 4°C with 10^7 magnetic beads (Epoxy

M270 Dynabeads, Invitrogen) covalently coated with purified rabbit IgG. Beads were washed three times with lysis buffer before resuspension in SDS-PAGE sample buffer. Samples were analyzed by SDS-PAGE and Western blotting.

^{3}H-uracil labeling

Cells were grown at 25°C in SC −URA to OD_{600} 0.6–0.9 and treated as described in the text. Five-milliliter aliquots were removed, supplemented with 50 μCi 5,6-^{3}H-uracil, and incubated for 20 min. Chase was performed for 20 min with a 100-fold excess cold uracil before total RNA was extracted and analyzed as described in the Supplemental Material.

Primer extension assays

Primer extension assays were performed as described previously (Claypool et al. 2004) with slight modifications. The protocol is detailed in the Supplemental Material.

Recombinant protein expression and purification–kinase assays

Purification of GST-Sch9$^{(3E/kd)}$ from yeast cells was performed as described previously (Urban et al. 2007), except that *SCH9^{3E}* was used as the active *SCH9* allele. Maf1 proteins were expressed in bacteria, affinity-purified using the pGEX6P1 system (Invitrogen), and assayed as substrates for GST-SCH9$^{3E/kd}$ as described previously (Urban et al. 2007).

Acknowledgments

We thank Tom Sturgill, Claudio DeVirgilio, David Shore, Howard Riezman, and R.A. and R.L. laboratory members for comments on draft manuscripts. We also thank the Functional Genomics Center Zurich for the generous support with MS resources. This project was funded in part by ETH Zurich, the Swiss National Science Foundation under grant no. 31000-10767, with federal (U.S.) funds from the National Heart, Lung, and Blood Institute, National Institutes of Health, under contract no. N01-HV-28179, and by SystemsX.ch, the Swiss initiative for systems biology. Work at the FGCZ was supported by the University Research Priority Program Systems Biology and Functional Genomics of the University of Zurich. R.A. was supported in part by a grant from F. Hoffmann-La Roche Ltd. (Basel, Switzerland) provided to the Competence Center for Systems Physiology and Metabolic Disease. B.B. was the recipient of a fellowship by the Boehringer Ingelheim Fonds. A.H. and A.U. were recipients of fellowships from the Novartis Foundation for Biology and Medicine and the Roche Research Foundation, respectively. R.L. is the recipient of a professorship from the Swiss National Science Foundation (PP00P3-110770/3100A0-108114) and receives generous support from the canton of Genève, the Leenaards Foundation, and the European Research Council (ERC-2007-StG 206173-TOR signaling).

References

Badis G, Chan ET, van Bakel H, Pena-Castillo L, Tillo D, Tsui K, Carlson CD, Gossett AJ, Hasinoff MJ, Warren CL, et al. 2008. A library of yeast transcription factor motifs reveals a widespread function for Rsc3 in targeting nucleosome exclusion at promoters. *Mol Cell* **32:** 878–887.

Beck T, Hall MN. 1999. The TOR signalling pathway controls nuclear localization of nutrient-regulated transcription factors. *Nature* **402:** 689–692.

Beugnet A, Tee AR, Taylor PM, Proud CG. 2003. Regulation of targets of mTOR (mammalian target of rapamycin) signalling by intracellular amino acid availability. *Biochem J* **372:** 555–566.

Bharucha N, Ma J, Dobry CJ, Lawson SK, Yang Z, Kumar A. 2008. Analysis of the yeast kinome reveals a network of regulated protein localization during filamentous growth. *Mol Biol Cell* **19:** 2708–2717.

Bodenmiller B, Malmstrom J, Gerrits B, Campbell D, Lam H, Schmidt A, Rinner O, Mueller LN, Shannon PT, Pedrioli PG, et al. 2007a. PhosphoPep–a phosphoproteome resource for systems biology research in Drosophila Kc167 cells. *Mol Syst Biol* **3:** 139.

Bodenmiller B, Mueller LN, Mueller M, Domon B, Aebersold R. 2007b. Reproducible isolation of distinct, overlapping segments of the phosphoproteome. *Nat Methods* **4:** 231–237.

Boguta M, Czerska K, Zoladek T. 1997. Mutation in a new gene MAF1 affects tRNA suppressor efficiency in Saccharomyces cerevisiae. *Gene* **185:** 291–296.

Brusniak MY, Bodenmiller B, Campbell D, Cooke K, Eddes J, Garbutt A, Lau H, Letarte S, Mueller LN, Sharma V, et al. 2008. Corra: Computational framework and tools for LC-MS discovery and targeted mass spectrometry-based proteomics. *BMC Bioinformatics* **9:** 542. doi: 10.1186/1471-2105-9-542.

Cherkasova VA, Hinnebusch AG. 2003. Translational control by TOR and TAP42 through dephosphorylation of eIF2α kinase GCN2. *Genes & Dev* **17:** 859–872.

Claypool JA, French SL, Johzuka K, Eliason K, Vu L, Dodd JA, Beyer AL, Nomura M. 2004. Tor pathway regulates Rrn3p-dependent recruitment of yeast RNA polymerase I to the promoter but does not participate in alteration of the number of active genes. *Mol Biol Cell* **15:** 946–956.

Cutler NS, Pan X, Heitman J, Cardenas ME. 2001. The TOR signal transduction cascade controls cellular differentiation in response to nutrients. *Mol Biol Cell* **12:** 4103–4113.

Dechant R, Peter M. 2008. Nutrient signals driving cell growth. *Curr Opin Cell Biol* **20:** 678–687.

Desai N, Lee J, Upadhya R, Chu Y, Moir RD, Willis IM. 2005. Two steps in Maf1-dependent repression of transcription by RNA polymerase III. *J Biol Chem* **280:** 6455–6462.

Desiere F, Deutsch EW, King NL, Nesvizhskii AI, Mallick P, Eng J, Chen S, Eddes J, Loevenich SN, Aebersold R. 2006. The PeptideAtlas project. *Nucleic Acids Res* **34:** D655–D658. doi: 10.1093/nar/gkj040.

De Virgilio C, Loewith R. 2006. Cell growth control: Little eukaryotes make big contributions. *Oncogene* **25:** 6392–6415.

Di Como CJ, Arndt KT. 1996. Nutrients, via the Tor proteins, stimulate the association of Tap42 with type 2A phosphatases. *Genes & Dev* **10:** 1904–1916.

Dilova I, Powers T. 2006. Accounting for strain-specific differences during RTG target gene regulation in *Saccharomyces cerevisiae*. *FEMS Yeast Res* **6:** 112–119.

Dilova I, Chen CY, Powers T. 2002. Mks1 in concert with TOR signaling negatively regulates RTG target gene expression in *S. cerevisiae*. *Curr Biol* **12:** 389–395.

Domon B, Aebersold R. 2006. Mass spectrometry and protein analysis. *Science* **312:** 212–217.

Duvel K, Santhanam A, Garrett S, Schneper L, Broach JR. 2003. Multiple roles of Tap42 in mediating rapamycin-induced transcriptional changes in yeast. *Mol Cell* **11:** 1467–1478.

Garrett S, Broach J. 1989. Loss of Ras activity in *Saccharomyces cerevisiae* is suppressed by disruptions of a new kinase gene, YAKI, whose product may act downstream of the cAMP-dependent protein kinase. *Genes & Dev* **3:** 1336–1348.

Geiduschek EP, Kassavetis GA. 2001. The RNA polymerase III transcription apparatus. *J Mol Biol* **310:** 1–26.

Jiang Y, Broach JR. 1999. Tor proteins and protein phosphatase 2A reciprocally regulate Tap42 in controlling cell growth in yeast. *EMBO J* **18:** 2782–2792.

Johnson SA, Dubeau L, Johnson DL. 2008. Enhanced RNA polymerase III-dependent transcription is required for oncogenic transformation. *J Biol Chem* **283:** 19184–19191.

Jorgensen P, Rupes I, Sharom JR, Schneper L, Broach JR, Tyers M. 2004. A dynamic transcriptional network communicates growth potential to ribosome synthesis and critical cell size. *Genes & Dev* **18:** 2491–2505.

Kaeberlein M, Powers RW III, Steffen KK, Westman EA, Hu D, Dang N, Kerr EO, Kirkland KT, Fields S, Kennedy BK. 2005. Regulation of yeast replicative life span by TOR and Sch9 in response to nutrients. *Science* **310:** 1193–1196.

Kamada Y, Funakoshi T, Shintani T, Nagano K, Ohsumi M, Ohsumi Y. 2000. Tor-mediated induction of autophagy via an Apg1 protein kinase complex. *J Cell Biol* **150:** 1507–1513.

Kim DH, Sarbassov DD, Ali SM, King JE, Latek RR, Erdjument-Bromage H, Tempst P, Sabatini DM. 2002. mTOR interacts with raptor to form a nutrient-sensitive complex that signals to the cell growth machinery. *Cell* **110:** 163–175.

Kim EM, Jang YK, Park SD. 2002. Phosphorylation of Rph1, a damage-responsive repressor of PHR1 in *Saccharomyces cerevisiae*, is dependent upon Rad53 kinase. *Nucleic Acids Res* **30:** 643–648.

Laferte A, Favry E, Sentenac A, Riva M, Carles C, Chedin S. 2006. The transcriptional activity of RNA polymerase I is a key determinant for the level of all ribosome components. *Genes & Dev* **20:** 2030–2040.

Lee J, Moir RD, Willis IM. 2009. Regulation of RNA polymerase III transcription involves SCH9-dependent and SCH9-independent branches of the target of rapamycin (TOR) pathway. *J Biol Chem* **284:** 12604–12608.

Lempiainen H, Uotila A, Urban J, Dohnal I, Ammerer G, Loewith R, Shore D. 2009. Sfp1 interaction with TORC1 and Mrs6 reveals feedback regulation on TOR signaling. *Mol Cell* **33:** 704–716.

Liko D, Slattery MG, Heideman W. 2007. Stb3 binds to ribosomal RNA processing element motifs that control transcriptional responses to growth in *Saccharomyces cerevisiae*. *J Biol Chem* **282:** 26623–26628.

Marshall L, Kenneth NS, White RJ. 2008. Elevated tRNA(iMet) synthesis can drive cell proliferation and oncogenic transformation. *Cell* **133:** 78–89.

Mayer C, Grummt I. 2006. Ribosome biogenesis and cell growth: mTOR coordinates transcription by all three classes of nuclear RNA polymerases. *Oncogene* **25:** 6384–6391.

Moir RD, Lee J, Haeusler RA, Desai N, Engelke DR, Willis IM. 2006. Protein kinase A regulates RNA polymerase III transcription through the nuclear localization of Maf1. *Proc Natl Acad Sci* **103:** 15044–15049.

Mueller LN, Rinner O, Schmidt A, Letarte S, Bodenmiller B, Brusniak MY, Vitek O, Aebersold R, Muller M. 2007. SuperHirn—a novel tool for high resolution LC-MS-based peptide/protein profiling. *Proteomics* **7:** 3470–3480.

Mueller LN, Brusniak MY, Mani DR, Aebersold R. 2008. An assessment of software solutions for the analysis of mass spectrometry based quantitative proteomics data. *J Proteome Res* **7:** 51–61.

Nanahoshi M, Nishiuma T, Tsujishita Y, Hara K, Inui S, Sakaguchi N, Yonezawa K. 1998. Regulation of protein phosphatase 2A catalytic activity by α4 protein and its yeast homolog Tap42. *Biochem Biophys Res Commun* **251:** 520–526.

Oficjalska-Pham D, Harismendy O, Smagowicz WJ, Gonzalez de Peredo A, Boguta M, Sentenac A, Lefebvre O. 2006. General repression of RNA polymerase III transcription is triggered by protein phosphatase type 2A-mediated dephosphorylation of Maf1. *Mol Cell* **22:** 623–632.

Olsen JV, Blagoev B, Gnad F, Macek B, Kumar C, Mortensen P, Mann M. 2006. Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* **127:** 635–648.

Pike BL, Yongkiettrakul S, Tsai MD, Heierhorst J. 2004. Mdt1, a novel Rad53 FHA1 domain-interacting protein, modulates DNA damage tolerance and G(2)/M cell cycle progression in *Saccharomyces cerevisiae*. *Mol Cell Biol* **24:** 2779–2788.

Pinkse MWH, Uitto PM, Hilhorst MJ, Ooms B, Heck AJR. 2004. Selective isolation at the femtomole level of phosphopeptides from proteolytic digests using 2D-nanoLC-ESI-MS/MS and titanium oxide precolumns. *Anal Chem* **76:** 3935–3943.

Planta RJ. 1997. Regulation of ribosome synthesis in yeast. *Yeast* **13:** 1505–1518.

Rinner O, Mueller LN, Hubalek M, Muller M, Gstaiger M, Aebersold R. 2007. An integrated mass spectrometric and computational framework for the analysis of protein interaction networks. *Nat Biotechnol* **25:** 345–352.

Roberts DN, Wilson B, Huff JT, Stewart AJ, Cairns BR. 2006. Dephosphorylation and genome-wide association of Maf1 with Pol III-transcribed genes during repression. *Mol Cell* **22:** 633–644.

Santhanam A, Hartley A, Duvel K, Broach JR, Garrett S. 2004. PP2A phosphatase activity is required for stress and Tor kinase regulation of yeast stress response factor Msn2p. *Eukaryot Cell* **3:** 1261–1271.

Schmidt A, Beck T, Koller A, Kunz J, Hall MN. 1998. The TOR nutrient signalling pathway phosphorylates NPR1 and inhibits turnover of the tryptophan permease. *EMBO J* **17:** 6924–6931.

Schmidt A, Gehlenborg N, Bodenmiller B, Mueller LN, Campbell D, Mueller M, Aebersold R, Domon B. 2008. An integrated, directed mass spectrometric approach for in-depth characterization of complex peptide mixtures. *Mol Cell Proteomics* **7:** 2138–2150.

Siebel CW, Feng L, Guthrie C, Fu XD. 1999. Conservation in budding yeast of a kinase specific for SR splicing factors. *Proc Natl Acad Sci* **96:** 5440–5445.

Steffen KK, MacKay VL, Kerr EO, Tsuchiya M, Hu D, Fox LA, Dang N, Johnston ED, Oakes JA, Tchao BN, et al. 2008. Yeast life span extension by depletion of 60s ribosomal subunits is mediated by Gcn4. *Cell* **133:** 292–302.

Upadhya R, Lee J, Willis IM. 2002. Maf1 is an essential mediator of diverse signals that repress RNA polymerase III transcription. *Mol Cell* **10:** 1489–1494.

Urban J, Soulard A, Huber A, Lippman S, Mukhopadhyay D, Deloche O, Wanke V, Anrather D, Ammerer G, Riezman H, et al. 2007. Sch9 is a major target of TORC1 in *Saccharomyces cerevisiae*. *Mol Cell* **26:** 663–674.

Venema J, Tollervey D. 1999. Ribosome synthesis in *Saccharomyces cerevisiae*. *Annu Rev Genet* **33:** 261–311.

Wanke V, Cameroni E, Uotila A, Piccolis M, Urban J, Loewith R, De Virgilio C. 2008. Caffeine extends yeast lifespan by targeting TORC1. *Mol Microbiol* **69:** 277–285.

Warner JR. 1999. The economics of ribosome biosynthesis in yeast. *Trends Biochem Sci* **24:** 437–440.

Wei M, Fabrizio P, Hu J, Ge H, Cheng C, Li L, Longo VD. 2008. Life span extension by calorie restriction depends on Rim15 and transcription factors downstream of Ras/PKA, Tor, and Sch9. *PLoS Genet* **4:** e13. doi: 10.1371/journal.pgen.0040013.

Willis IM, Moir RD. 2007. Integration of nutritional and stress signaling pathways by Maf1. *Trends Biochem Sci* **32:** 51–53.

Willis IM, Desai N, Upadhya R. 2004. Signaling repression of transcription by RNA polymerase III in yeast. *Prog Nucleic Acid Res Mol Biol* **77:** 323–353.

Woiwode A, Johnson SA, Zhong S, Zhang C, Roeder RG, Teichmann M, Johnson DL. 2008. PTEN represses RNA polymerase III-dependent transcription by targeting the TFIIIB complex. *Mol Cell Biol* **28:** 4204–4214.

Wullschleger S, Loewith R, Hall MN. 2006. TOR signaling in growth and metabolism. *Cell* **124:** 471–484.

Zaragoza D, Ghavidel A, Heitman J, Schultz MC. 1998. Rapamycin induces the G0 program of transcriptional repression in yeast by interfering with the TOR signaling pathway. *Mol Cell Biol* **18:** 4463–4470.

Zhang C, Comai L, Johnson DL. 2005. PTEN represses RNA polymerase I transcription by disrupting the SL1 complex. *Mol Cell Biol* **25:** 6899–6911.

Zhou HL, Watts JD, Aebersold R. 2001. A systematic approach to the analysis of protein phosphorylation. *Nat Biotechnol* **19:** 375–378.

Zhu C, Byers K, McCord R, Shi Z, Berger M, Newburger D, Saulrieta K, Smith Z, Shah M, Radhakrishnan M, et al. 2009. High-resolution DNA binding specificity analysis of yeast transcription factors. *Genome Res* **19:** 556–566.

C

THE DUAL SPECIFICITY KINASE DYRK3 COUPLES STRESS GRANULE CONDENSATION/DISSOLUTION TO MTORC1 SIGNALING

For this work I processed the LC–MS/MS data generated using the clustering algorithm for LC–MS/MS alignments described in chapter 7.

The manuscript included below is currently under second review.

The dual specificity kinase DYRK3 couples stress granule condensation/dissolution to mTORC1 signaling.

Frank Wippich[1,2], Bernd Bodenmiller[1,2], Maria Trajkovska[1], Stefanie Wanka[1], Ruedi Aebersold[2,3] and Lucas Pelkmans[1,a]

[a] Correspondence: lucas.pelkmans@imls.uzh.ch (L.P.)

[1] Institute of Molecular Life Sciences, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland.
[2] Institute of Molecular Systems Biology, ETH Zurich, Wolfgang-Pauli-Strasse 16, CH-8093 Zurich, Switzerland.
[3] Faculty of Sciences, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland.

Summary

Cytosolic compartmentalization through liquid-liquid unmixing, such as the formation of stress granules during cellular stress, is involved in many cellular processes and may be coupled to signal transduction, but specific molecular regulators of this are unknown. Here, we reveal that the dual specificity kinase DYRK3 is able to condense P granule-like structures and to prevent stress granule dissolution via its N-terminal domain when inhibited. DYRK3 couples this to mTORC1 signaling by sequestering mTORC1 on stress granules when inactive and by phosphorylating PRAS40 when active. By showing that DYRK3 dynamically regulates its own partitioning between SGs and the cytosol through its kinase activity, our work suggests a novel cyclic partitioning mechanism that couples compartmentalization through liquid phase transitions with cellular signaling.

Introduction

The nucleus and cytosol of eukaryotic cells contain numerous nonmembrane-bound compartments that consist of many proteins involved in complex reactions. Well known amongst these are different types of RNA granules, microscopically visible accumulations of messenger ribonucleo-protein (mRNP) (Anderson and Kedersha, 2009; Eulalio et al., 2007). A complex repertoire of mRNP-associated proteins determine whether mRNPs complexes remain silent, become translationally active, or are degraded (Buchan and Parker, 2009). Especially during stressful conditions such as heat, oxidative stress, virus infection, osmotic stress and UV irradiation, mRNPs accumulate into large granules, so called stress granules (SGs), where they are translationally silenced (Anderson and Kedersha, 2008).

Recently, it has become clear that the accumulation of RNA granules is reminiscent of concentration-dependent liquid-liquid de-mixing of mRNA and proteins with low complexity domains (Hyman and Simons, 2012; Kato et al., 2012; Weber and Brangwynne, 2012). While this may be a unifying principle of compartmentalization without membranes (Brangwynne et al., 2009), many fundamental questions remain unanswered. For instance, are there specific molecular regulators of this type of compartmentalization, and is it utilized to control signal transduction, analogous to membrane-bound compartments? These questions become apparent during cellular stress, when cells have to coordinate SG condensation and dissolution with the control of signaling

pathways that initiate mRNA translation (Buchan and Parker, 2009), amongst which the mechanistic target of rapamycin (mTOR) signaling pathway takes a prominent role (Loewith and Hall, 2011; Ma and Blenis, 2009; Sengupta et al., 2010; Zoncu et al., 2010). Interestingly, in *S. cerevisiae* the mTOR complex 1 (mTORC1) partitions in heat-induced SGs, suggesting a coupling of SG condensation and dissolution with mTORC1 signaling (Takahara and Maeda, 2012). How this is coupled is however unclear.

Here, we identify the dual specificity tyrosine-phosphorylation-regulated kinase 3 (DYRK3) as a protein with the ability to condense P granule-like speckles in the cytosol and to prevent SG dissolution via its N-terminal domain when it is in a kinase-inactive form. DYRK3 couples this to the control of mTORC1 signaling by sequestering mTORC1 on SGs when inactive, and by phosphorylating PRAS40 in the cytosol, a negative regulator of mTORC1 (Sancak et al., 2007; Vander Haar et al., 2007), when active. Since DYRK3 dynamically regulates its own partitioning between SGs and the cytosol through its kinase activity, this suggests a novel cyclic partitioning mechanism that couples compartmentalization through liquid phase transitions with cellular signaling.

RESULTS

An image-based screen for chemical compound inhibitors that delay stress granule disassembly

To study the regulation of SG dissolution, we screened a custom library of 246 small compound kinase inhibitors in their ability to prolong the presence of SGs in HeLa cells during recovery from arsenite-induced oxidative stress. We first exposed cells to arsenite for 45 minutes, after which we allowed recovery for 240 minutes in the presence of small compound inhibitors at 2 different concentrations (Figure 1A). To monitor the occurrence of SGs, we used immuno-fluorescence staining against polyadenylate-binding protein 1 (PABP1), which, upon stress, gets recruited to SGs, resulting in bright and easily detectable granules (Kedersha et al., 1999). Quantification of the fraction of cells containing SGs was performed with automated image analysis and machine learning.

Compounds that were able to block SG dissolution at least more than 3 standard deviations above the mean of all treatments were considered as significant hits (z-score ≥ 3; red dots). Based on this, we identified two compounds at 1 μM concentration, GSK-626616 and Jak3 inhibitor IV as most potent in blocking SG dissolution (Figure 1B, left side). At 10 μM concentration, we identified the compounds Harmine, Wee1/Chk1 Inhibitor, CDK4 Inhibitor III, CDK2 Inhibitor IV, PKR inhibitor negative control and GSK-626616 as most potent in blocking SG disassembly (Figure 1B, right side).

To validate the candidate compounds found in the initial screen, we performed a second stress recovery assay using time-lapse imaging of RDG3 cells expressing GFP-tagged Ras GTPase-activating protein-binding protein 1 (G3BP1). These cells were previously used in a RNA interference (RNAi) screen to identify genes involved in mRNA granule condensation (Ohn et al., 2008). Quantitative analysis showed that oxidative stress for 45 minutes caused a rapid increase of SG-positive cells, followed by a rapid decline in the subsequent recovery phase after arsenite washout (Figure 1C and Suppl. Movie S1). From the 7 compounds identified in the initial screen, we could confirm 5 compounds, GSK-626616, Harmine, CDK2 Inhibitor IV, CDK4 Inhibitor III and Wee1/Chk1 Inhibitor, to reduce the decline rate of SG-positive cells during stress recovery in a second cell line using another SG marker.

We further tested which of these compounds can induce SG condensation in the absence of stress. Treatment of cells for 240 min with these compounds in the absence of any stress revealed that CDK2 Inhibitor IV and to a lesser extent Wee1/Chk1 Inhibitor are able to induce SGs (Figure 1D). CDK4 inhibitor III, GSK-626616, and Harmine did not cause significant SG formation in HeLa cells during 240 min of treatment.

Finally, we analyzed whether the compounds have an effect on the phosphorylation of eukaryotic initiation factor 2 alpha (eIF2α) at Ser51, which is induced by arsenite treatment and causes a translational arrest and rapid appearance of SGs (Anderson and Kedersha, 2002). Interestingly, none of the compounds could trigger the phosphorylation of eIF2α on its own, but the

Wee1/Chk1 Inhibitor prevented dephosphorylation of eIF2α after recovery from oxidative stress (Figure 1D).

Taken together, these experiments revealed that GSK-626616, Harmine and CDK4 Inhibitor III are able to reduce the rate of SG dissolution after oxidative stress without evoking *de novo* condensation of SGs, and without affecting phosphorylation and de-phosphorylation of eIF2α. Clearly, GSK-626616 was the most potent compound, being able to delay SG disassembly at 1 μM. Furthermore, Harmine was present in top 4% of all compounds tested at low concentrations (Figure 1B, left side), and was the top hit at 10 μM. Interestingly, both GSK-626616 and Harmine are known inhibitors of DYRK family kinases (Erickson-Miller et al., 2007; Gockler et al., 2009). GSK-626616 has a reported IC_{50} for DYRK3 of 0.7 nM, but overall kinase specificity is unclear. To obtain a kinome-wide view on the kinase specificity of GSK-626616, we profiled the inhibitory effect of GSK-626616 on 451 kinases *in vitro* (Karaman et al., 2008). At 0.1 μM, GSK-626616 primarily inhibits DYRK family kinases (Suppl. Figure S1A and Table 2) with a good selectivity (S-Score(35) = 0.07), comparable to well studied and highly specific kinase inhibitors (Karaman et al., 2008). At 1 μM, GSK-626616 displays some off-target effects *in vitro*, but still has reasonable selectivity against DYRK family kinases (S-Score(35) = 0.12). Harmine was previously tested on 67 kinases *in vitro*, and also seems to have a specific potential in inhibiting DYRK family kinase at 1 μM (Bain et al., 2007), in particular of DYRK1A, DYRK2, and DYRK3 (Gockler et al., 2009).

The compound screen and *in vitro* kinase profiling revealed that specific inhibitors of DYRK family kinases delay the dissolution of SGs during recovery from stress in a manner that does not involve classical SG condensation pathways via eIF2α.

DYRK3 can condense P granule-like speckles and partitions in stress granules

Consistent with a previous classification of DYRK family kinases by their subcellular localization, we found that GFP-DYRK1A and GFP-DYRK1B localized to the nucleus (Suppl. Figure S2A). Interestingly, GFP-DYRK3 localized predominantly on distinct speckles distributed throughout the cytoplasm of the cell (Figure 2A). This subcellular localization of DYRK3 was not seen for DYRK2 and DYRK4, the other class 2 DYRK family members (Suppl. Figure S2A).

However, at low levels of expression, GFP-DYRK3 distributed homogeneous throughout the cytoplasm and did not aggregate in speckles (Figure 2B). When we quantified this aggregating effect in a large number of cells, we discovered that at a defined expression-level, DYRK3 suddenly condenses in speckles (Figure 2B). Furthermore, following GFP-DYRK3 speckles by time-lapse microscopy revealed that the speckles are dynamic and can merge in a liquid-like manner (Suppl. Figure S2B and Movie S2). These observations are reminiscent of P granule behavior in *C. elegans* (Brangwynne et al., 2009). In vertebrates as well as in *C. elegans*, GW182 is a scaffold protein of mRNA

processing-bodies (P-bodies), sites of mRNA decay and storage (Eulalio et al., 2007; Eystathioy et al., 2003). Indeed, when co-expressed, DYRK3 and GFP-GW182 co-localize in numerous granules that were somewhat larger than formed by GFP-DYRK3 expression alone (Figure 2C).

Importantly, we observed that during oxidative and osmotic stress, endogenous DYRK3 and GFP-DYRK3 localize to SGs (Figure 2D and 2E). Furthermore, DEAD box p54 protein 6 (DDX6)-positive P-bodies, often in close proximity to SGs, were found docked on GFP-DYRK3-positive granules (Suppl. Figure S2C). Moreover, endogenous DYRK3 localized to SGs, present 240 minutes after stress and retained by the treatment with DYRK inhibitors after stress (Figure 2F and S2D).

Thus, DYRK3 has the potential to condense granules in the cytosol of human cells on which the mRNA-binding protein GW182 can be recruited, and DYRK3 localizes to SGs during oxidative and osmotic stress.

The N-terminal domain of DYRK3 is required for stress granule localization and induces stress granules when DYRK3 kinase activity is compromised

We next wondered how inhibition of DYRK3 prevents SG dissolution. Surprisingly, we observed that RNAi-mediated depletion of DYRK3 or of the other DYRKs (not shown) did not disturb the dissolution of SGs after oxidative stress (Figure 3A). Instead, we discovered that DYRK3 depletion, but not that of the other DYRKs (not shown) reduced the block in SG dissolution caused

by GSK-626616 treatment. This suggests that inhibited DYRK3 is in a state that specifically prevents the dissolution of SGs. In support of this, we observed that expression of the kinase-deficient point-mutant DYRK3-K218M is sufficient to cause the appearance of large cytoplasmic structures positive for mRNA granule markers, on which GFP-DYRK3-K218M accumulated, even in the absence of stress (Figure 3B). To test if a specific domain of DYRK3 mediates the partitioning in RNA granules, we generated a series of DYRK3 truncations (Figure 3C). Expression of DYRK3-NT, which consists of the N-terminal residues 1-188 and contains a predicted low-complexity sequence, but which excludes the kinase domain and C-terminal end, induced the appearance of large granules at high expression levels, which, similar to SGs, had P-bodies in close proximity and stained positive for PABP1 (Figure 3D and 3E). Conversely, overexpression of DYRK3 without the N-terminal domain (DYRK3-ΔNT) did not induce cytoplasmic granules, and did not partition in SGs induced by oxidative stress (Figure 3F).

Thus, inhibition of DYRK3 affects SG dissolution through a specific state of DYRK3 when its kinase activity is compromised. This state depends on the N-terminal domain of DYRK3, which, when expressed alone, is able to induce the appearance of SGs in the absence of stress.

Inhibition of DYRKs affects the phosphorylation of proteins that bind mRNA, partition in stress granules, and are downstream of mTORC1

To obtain insight into how cells couple SG condensation and dissolution to signal transduction, we next studied which signaling pathways are affected by the inhibition of DYRKs. To reveal this, we studied changes in the phospho-proteome of cells after GSK-626616 treatment, using a quantitative label-free phospho-proteomic approach (Bodenmiller et al., 2010; Huber et al., 2009) (Figure 4A). For two different time-points, 30 minutes and 12 hours respectively, we monitored the abundance of phospho-peptides isolated with titanium dioxide-based affinity chromatography and captured by liquid chromatography tandem mass spectrometry (LC-MS), reflecting the phosphorylation state of peptides (Bodenmiller et al., 2007). Of the overall 1,194 peptides identified, only peptides that were present in three independent replicate experiments, and which were significantly enriched or depleted in the GSK-626616-treated compared to control samples (t-test P-value: <0.1), where taken into account (Table 3). These strict criteria yielded a total of 44 regulated phospho-peptides: 26 regulated phospho-peptides after short-term treatment and 18 regulated phospho-peptides after long-term treatment with GSK-626616 (Figure 4B). 6 phospho-peptides were found significantly enriched or depleted at both time-points.

20 out of the 24 and 7 out of the 15 corresponding proteins that were affected by short- and long-term treatment with GSK-626616 respectively, were recently reported to associate with mRNA or mRNA granules (Baltz et al., 2012; Castello et al., 2012; Elvira et al., 2006; Kato et al., 2012). In addition, mTORC1 signaling regulates 5 of the affected proteins in each treatment (Hsu et al., 2011; Yu et al., 2011), such as the eukaryotic translation initiation factor

4E-binding protein 1 (eIF4E-BP1) at threonine 37 and 46, and the tumor suppressor protein programmed cell death 4 (PDCD4) at serine 457. Both eIF4E-BP1 and PDCD4 are translational repressors known to act downstream of the mTORC1-ribosomal protein S6 kinase 1 (S6K1) pathway (Brunn et al., 1997; Burnett et al., 1998). Similar changes in the phosphorylation of eIF4E-BP1 and PDCD4 were detected by immunoblotting using phospho-specific antibodies, confirming our findings (Suppl. Figure S4A). Thus, the DYRK kinase inhibitor GSK-626616 affects, amongst others, the phosphorylation of mRNA-associated proteins and proteins downstream of mTORC1 signaling.

DYRK3 is required for mTORC1 activity

To validate the impact of inhibitors targeting DYRK family kinases on mTORC1 signaling, we evaluated the phosphorylation state of Thr389 of S6K1, a phosphorylation site that is considered a direct and appropriate readout for mTORC1 activity (Burnett et al., 1998). GSK-626616 treatment abolished the phosphorylation of S6K1 at Thr389 in non-stimulated HeLa cells, indicating a reduction in basal mTORC1 activity (Figure 4C). A variety of stimuli, such as epidermal growth factor (EGF) and insulin addition, can increase mTORC1 activity above basal levels (Zoncu et al., 2010). GSK-626616 treatment also reduced the phosphorylation of S6K1 at Thr389 in EGF- as well as insulin-stimulated HeLa cells, showing that mTORC1 activity is impaired (Figure 4D). The extent by which low concentrations of GSK-626616 inhibited S6K1 phosphorylation in cells was similar to the effect of a rapamycin derivative (RAD-001), a well-known inhibitor of mTORC1 (Boulay

et al., 2004). Also Harmine treatment was able to interfere with EGF-stimulated S6K1 phosphorylation, although to a lesser degree (Suppl. Figure S4B). Moreover, GSK-626616 treatment interfered with basal mTORC1 activity, as well as EGF- or insulin- stimulated mTORC1 activity in 3 additional unrelated mammalian cell lines, indicating a general mechanism that is not cell line-specific (Suppl. Figure S4C). As a consequence, treatment of cells with GSK-626616 leads to a reduction in protein synthesis (Suppl. Figure S4G).

The mTORC1 pathway integrates many different signaling inputs, and the DYRK kinase inhibitors may affect several of those. However, we observed no effect of GSK-626616 treatment on the phosphorylation status of tuberin (TSC2) upon stimulation of cells with either EGF or insulin (Figure 4E). In addition, we observed no effect on the phosphorylation of AKT or ERK1/2 (Suppl. Figure S4D). We also found that the PI-3 kinase inhibitor Wortmannin had an additive inhibitory effect on mTORC1 activity in cells treated with GSK-626616 (Suppl. Figure S4E), further indicating that the DYRK inhibitor affects mTORC1 via a parallel pathway. To assess whether the DYRK inhibitors could affect mTORC1 activity via an off-target effect and not via inhibiting DYRKs, we considered the few *in vitro* off-targets that these inhibitors have (Suppl. Figure S4F). Two possible kinases, namely ERK8 and RSK3/4, that are weakly inhibited by GSK-626616 *in vitro*, might act upstream of mTORC1. ERK8 however, is not inhibited by Harmine (Bain et al., 2007), and both ERK and RSK family kinases signal to mTORC1 via TSC2, which phosphorylation status does not change during GSK-626616 treatment (Figure 4E). Finally,

siRNA- and shRNA-mediated depletion of DYRK3 (but not other DYRK family kinases, data not shown) also resulted in a reduced S6K1 phosphorylation (Figure 4F and 4G). This further supports a specific effect of the inhibitors, and indicates that the inhibitory effect on mTORC1 acts mainly through DYRK3 in a manner that is independent of the PI3K/AKT and ERK pathways, and does not involve TSC2.

Kinase-inactive DYRK3 inhibits mTORC1 by preventing dissociation from SGs.

One possibility by which inhibition of DYRK3 may result in an inhibition of mTORC1 signaling activity is by causing a partitioning of mTORC1 into SGs. That mTORC1 partitions in SGs was recently shown in *S. cerevisiae* (Takahara and Maeda, 2012). Interestingly, we found that in human cells, both mTOR and the mTORC1-specific component RAPTOR are recruited to SGs induced by either osmotic or oxidative stress, and were retained on SGs by GSK-626616 treatment (Figure 5A and S5A).

While examining mTORC1 activity during and after stressful conditions, we observed a reduction in mTORC1 activity during stress, as reported previously (Inoki et al., 2003), and a hyper-activation of mTORC1 after recovery (Figure 5B). The reactivation of mTORC1 was blocked by the presence of GSK-626616 during recovery. To test whether this is a consequence of prolonged mTORC1 sequestration on SGs, we made use of the SG-dissolving property of cycloheximide (CHX), which rapidly increases the rate of SG disassembly

and accelerates mTORC1 reactivation (Buchan and Parker, 2009; Takahara and Maeda, 2012). When we added CHX to cells during recovery from stress, both SG disassembly and reactivation of mTORC1 were enhanced (Figure 5B and 5C). However, addition of CHX to cells during recovery from stress in the presence of GSK-626616, led to rapid disassembly of SGs, but not to a full reactivation of mTORC1. This indicates that DYRK inhibitors reduce mTORC1 signaling by blocking SG dissolution as well as by a second mechanism, independent of mTORC1 partitioning in SGs.

DYRK3 also controls mTORC1 by direct phosphorylation of PRAS40

To find this second mechanism, we used a microarray for kinase substrate identification consisting of more than 9,000 different human recombinant proteins. We identified 26 candidate proteins to be directly phosphorylated *in vitro* by wild type DYRK3 but not in by kinase-deficient DYRK3-K218M (Suppl. Figure S6A and Table 4). The second most strongly phosphorylated protein we identified the proline-rich AKT substrate of 40 kDa (PRAS40), also known as AKT1 substrate 1, which has been shown to interact directly with the mTORC1 complex and to negatively regulate mTOR kinase activity (Sancak et al., 2007; Vander Haar et al., 2007). Other proteins on the microarray that are known to regulate mTORC1 activity were not directly phosphorylated by DYRK3 (Table 4).

Phosphorylation of PRAS40 by AKT1 at Thr246 has been shown to release it from the mTORC1 complex, thereby abolishing the inhibitory effect of

PRAS40 on mTOR, which allows further activation (Sancak et al., 2007; Vander Haar et al., 2007). Using recombinant DYRK3 and PRAS40, we found that DYRK3 directly phosphorylates Thr246 of PRAS40 in an *in vitro* kinase assay (Figure 6A). Addition of GSK-626616 to this assay blocked the phosphorylation reaction in a dose-dependent manner (Suppl. Figure S6B). Moreover, in cultured cells, GSK-626616 treatment reduced the phosphorylation of PRAS40 at Thr246 in response to EGF treatment (Figure 6B). Using co-immunoprecipitation experiments, we observed that GSK-626616 increases the fraction of bound PRAS40 to mTORC1 (Figure 6C), which is known to interfere with binding and activation of mTOR substrates (Vander Haar et al., 2007). Thus, DYRK3 directly phosphorylates PRAS40 at Thr246, a phosphorylation site responsible for regulation of PRAS40, resulting in decreased binding of PRAS40 to mTORC1, allowing activation of mTORC1 signaling in unstressed cells and reactivation of mTORC1 during stress recovery.

DYRK3 regulates its own partitioning between SGs and the cytosol in a cyclic manner through its kinase activity

Finally, we studied the dynamics of how DYRK3 couples SG partitioning with mTORC1 regulation. To reveal this, we performed fluorescence recovery after photo-bleaching (FRAP) experiments. In unstressed cells expressing GFP-DYRK3 and displaying cytoplasmic speckles, we observed that photo-bleached speckles quickly recovered their fluorescence signal to initial levels

within 150 seconds (Figure 7A). Interestingly, treatment of GSK-626616 during stress recovery resulted in an entrapment of 55% of GFP-DYRK3 on granules, which did not exchange with unbleached cytosolic DYRK3-GFP. Similarly, when we photo-bleached the cytoplasmic aggregates induced by the kinase-deficient mutant of DYRK3 (GFP-DYRK3-K218M), we observed that 69% of the signal could not be exchanged with unbleached cytosolic signal. Thus, DYRK3 displays dynamic cycles of association to and dissociation from SGs. Association requires the N-terminal domain (deletion of this domain abolishes partitioning in SGs, see Figure 3F), while dissociation requires kinase activity.

Based on this, we propose the following mechanism by which DYRK3 couples SG condensation and dissolution with mTORC1 signaling. When SGs condense, such as during stress, DYRK3 will partition in SGs via its N-terminal domain where it contributes to preventing SG dissolution, leading to partitioning of the mTORC1 complex in SGs, preventing it from signaling to downstream effectors. To dissolve SGs, the kinase activity of DYRK3 is required, leading to partitioning of the mTORC1 complex in the cytosol, where DYRK3 allows mTORC1 to be activated by phosphorylating PRAS40. Thus, DYRK3 represents a new type of regulator that dynamically couples phase transition-mediated compartmentalization to signal transduction via its kinase activity.

Discussion

In this study we identify chemical compounds targeting DYRK family kinases as inhibitors of SG dissolution, and show that these compounds act mainly via DYRK3. We reveal that DYRK3 has the potential to condense P-granule-like structures in the cytosol, and localizes to SGs after stress. Not the absence of DYRK3, but kinase-inhibited DYRK3 prevents SG dissolution via its N-terminal domain. This domain is, when expressed alone or as part of a kinase-deficient mutant of DYRK3, able to induce the appearance of SG-like structures even in the absence of stress. We also show that inhibition of DYRKs affects the phosphorylation status of a number of mRNA-binding proteins and of proteins downstream of mTORC1 signaling. We explain this by demonstrating that DYRK3 blocks mTORC1 signaling by keeping mTORC1 partitioned in SGs when inactive, and phosphorylates the mTORC1 inhibitor PRAS40 when active, which reduces the binding of PRAS40 to mTORC1, allowing subsequent activation.

The condensation of SGs via liquid-liquid de-mixing depends on low complexity domains in proteins (Hyman and Simons, 2012; Kato et al., 2012; Weber and Brangwynne, 2012). Also the N-terminal domain of DYRK3 contains a predicted low complexity sequence, which may explain how it allows partitioning into stress granules and contribute to liquid-liquid unmixing. This is however not a constitutive property of the DYRK3 protein, but is regulated by its own kinase activity in a cyclic manner.

Such a kinase-activity-dependent cycle provides an RNA granule sensing mechanism. When SGs condense during stress, DYRK3 senses this through its ability to cycle between SGs and the cytosol, while mTORC1 partitions in SGs, which prevent signaling to downstream effectors. When stress signals are gone, the kinase activity of DYRK3 will dissolve SGs by 'inactivating' the SG condensation property of its N-terminal domain, and possibly by phosphorylating mRNA-binding proteins, and will allow activation of mTORC1 by preventing binding to PRAS40. In addition, DYRK3 might be a sensor of RNA granules that are formed under conditions different than stress, such as during cell division, cell polarization and cell differentiation, and to link their appearance to the control of mTORC1 signaling or other signal transduction pathways. We expect that following up on the numerous other proteins which phosphorylation status is affected by inhibition of DYRKs, or which can be phosphorylated by DYRK3 *in vitro* will provide further detail into this.

Several lines of evidence suggest that the properties of DYRK3 that we have uncovered here is a conserved mechanistic principle for DYRK family members. DYRKs have been functionally linked to the cellular stress response in organisms across the eukaryotic kingdom, ranging from nutrient starvation, osmotic stress, irradiation, and genotoxic stress (Aranda et al., 2010; Moriya et al., 2001; Seifert and Clarke, 2009; Taira et al., 2007; Taminato et al., 2002; Zhang et al., 2005). Furthermore, DYRKs are constitutively active kinases, believed to be regulated by changing their subcellular localization (Aranda et al., 2010). The dynamic cycling mechanism of DYRK3 between SGs and the cytosol shares striking resemblance with that

of Pom1, a DYRK kinase in *S. pombe* (Hachet et al., 2011). Pom1 cycles between a membrane-associated state, driven by auto-phosphorylation, and a cytosolic state, driven by phosphatase-mediated dephosphorylation of its N-terminus. Our results support a similar mechanism for DYRK3. An unphosphorylated N-terminal low complexity domain might partition in SGs, while auto-phosphorylation of this domain might abolish this property, keeping it in the cytosol. This would explain how kinase-inactive DYRK3 becomes trapped in SGs, as its N-terminal domain now constitutively partitions into SGs. However, an autophosphorylation event in the N-terminal domain of DYRK3 remains to be identified, as well as a phosphatase that would dephosphorylate it.

At least two other DYRK family members have been functionally linked to intracellular granules or speckles. MBK-2, a *C. elegans* DYRK, has been observed in speckles in the cytoplasm (Stitzel et al., 2006), and is essential for the asymmetric distribution of P granules in the first cell division of the *C. elegans* embryo (Pang et al., 2004). This process depends on specifically lowering the condensation point for P granules at one site of the dividing embryo (Brangwynne et al., 2009). Analogous to the properties of DYRK3, MBK-2 might thus be involved in P granule condensation in *C. elegans*. Furthermore, mammalian DYRK1A accumulates in nuclear splicing speckles (see also Suppl. Figure S2A) and is capable of dissolving speckles, depending on its kinase activity (Alvarez et al., 2003). We also observed that DYRK2, most closely related to DYRK3, localized to SGs in its inhibited state where it is partly responsible for blocking SG dissolution (data not shown).

This suggests that DYRKs are a family of kinases that couple compartmentalization through liquid phase transitions with cellular signaling through a novel kinase-dependent cyclic partitioning mechanism. Future work combining structural studies and detailed mechanistic analysis of DYRK cycling in and out liquid-unmixed compartments with the identification of possible upstream factors that influence this cycle will provide further insight into how this is utilized in the regulation of various cellular processes.

Experimental Procedures

Materials

Antibodies were obtained from the following sources: antibody to DYRK3 was from Aviva Systems Biology, PABP1, c-myc and mTOR were from Santa Cruz Biotechnology; antibodies to S6K1, pThr389-S6K1, pThr308-Akt, eIF4E-BP1, pThr37/46-eIF4-BP1, PRAS40, pThr246-PRAS40, eIF2α, pSer51-eIF2α, TSC2 and pThr1462-TSC2 were from Cell Signaling Technology; antibody to pThr202/pTyr204-ERK1/ pThr184/pTyr186-ERK2 was from BD Transduction; antibody to DDX6 was from Bethyl Laboratories Inc.; antibodies to beta-Tubulin, PDCD4 and pSer457-PDCD4 were from Abcam; antibody to GST was from Sigma Aldrich, secondary antibodies conjugated to Alexa Fluor's were from Molecular Probes, secondary antibodies conjugated to were from Dianova.

Lipofectamine 2000, DMEM and Glutamax were from Invitrogen; Human recombinant insulin, DMSO, sodium-arsenite, Harmine, sorbitol, Triton X-100, DAPI, L-glutathione reduced, myelin basic protein (MBP) from bovine brain, Albumin from bovine serum (BSA) and fetal bovine serum were from Sigma Aldrich; Paraformaldehyde (PFA) solution was from Electron Microscopy sciences; γ^{33}P-ATP was from PerkinElmer; RAD-001 was from Selleck Chemicals; human recombinant EGF was from Millipore; human recombinant PRAS40 was from Biomol; GSK-626616 (PubChem CID: 15981157) was obtained from Andreas Tako (Institute for Organic Chemistry, University of

Innsbruck, Austria). Remaining compounds were from EMD Calbiochem (Cat. No. 539743 and 539746).

Cell Lines and Tissue Culture

HeLa cells were from Marino Zerial (MPI-CBG, Dresden), RDG3 were from Paul J. Anderson (Division of Rheumatology, Immunology and Allergy, Brigham and Women's Hospital, Boston), A431 and MEF were from ATCC (Molsheim Cedex). All cells were maintained at 37°C under 5% CO_2 in DMEM supplemented with 10% fetal bovine serum FBS and Glutamax. Prior treatments, cells were washed and serum deprived for 14 hours.

Stress granule assembly and disassembly assay

Cells were grown in 384-well plates (for initial screen, Greiner) or 8-well Lab-Tek chambers (for life cell imaging, VWR) in DMEM containing 10% FBS. 3 hours prior to the experiment, the medium was exchanged to DMEM without FBS. Cells were treated for 45 minutes with 0.5mM arsenite, washed twice with Medium and allowed to recover for 240 minutes in DMEM suplememted with compound inhibitors at indicated concentrations. Treatment was stopped by addition of PFA to a final concentration of 4%. SGs were stained by immuno-fluorescence against PABP1 and images were acquired using on an ImageXpress Micro microscope (Molecular Devices). Life cell imaging was carried out on a VisiScope Confocal Cell Explorer. Automatic cell detection was performed using CellProfiler (Carpenter et al., 2006), and features describing intensities and textures were used for cell classification by support vector machine learning using CellClassifier (Ramo et al., 2009).

In vitro kinase profiling

GSK-626616 compound was profiled against 451 kinases in an high-throughput active site-directed competition binding assay (KINOMEscan™, DiscoveRX) at concentrations of 0.1 μM and 1 μM. Data was visualized using TREEspot™ Compound Profile Visualization Tool.

Immuno-fluorescence and confocal microscopy

HeLa cells were grown on coverslips until the appropriate assay was performed. Cells were fixed by adding PFA to a final concentration of 4% and permeabilized with 0.1% Triton X-100. Unspecific binding was reduced by incubation in 1% BSA for 30 minutes. Cells were incubated with primary antibodies for 1 hour, followed by incubation with labeled secondary antibody for 1 hour. Nuclei were stained using DAPI. Imaging was carried on: Leica SP2-FCS confocal microscope equipped with a 63x 1.4-0.6NA DIC, Oil, HCX Plan-Apochromat objective; Leica SP5 Mid UV-VIS equipped with a 63x 1.4-0.6NA DIC, Oil, HCX Plan-Apochromat; Zeiss LSM710 equipped with a 63x 1.4NA Oil DIC Plan-Apochromat objective. Images were processed using ImageJ (http://rsb.info.nih.gov/ij/).

Phospho-proteomics

Cells were exposed to GSK-626616 or DMSO only in DMEM without FBS for 30 minutes and 12 hours respectively, washed with PBS and harvested in lysis-buffer (150 mM NaCl, 10 mM Tris pH 7.2, 10 mM EDTA, 1 mM Na_3VO_4, 200 nM Ocadaic acid, 20 nM Calyculin A, EDTA free protease inhibitor (Roche), 1 mM PMSF and 0.1% *Rapi*Gest™ (Waters)). Three replicas were independently collected and processed separately. The isolation of phospho-

peptides using titanium-dioxid, mass spectrometry analysis and quantitative analysis was performed as described previously (Huber et al., 2009). Candidate protein classes were determined using PANTHER database (www.pantherdb.org).

In vitro kinase substrate identification

ProtoArray® Human Protein Microarrays (PAH052406, Invitrogen) were blocked with 1% BSA in PBS for 3 hours at 4°C. Kinase Buffer (100 mM MOPS pH 7.2, 1% Nonidet P40, 100 mM NaCl, 10 mg/ml BSA, 5 mM $MgCl_2$, 5 mM $MnCl_2$), supplemented with 33 nM γ^{33}P-ATP and 0.5 μg recombinant GST-DYRK3 or GST-DYRK3-K218M were added per array and incubated for 1h. Arrays were washed with 0.5 % SDS, dried and exposed to Amersham Hyperfilm (GE Healthcare) for 3 hours. Analysis was carried out with ProtoArray® Prospector (Invitrogen).

RNAi

Cells were reverse transfected using Lipofectamine 2000 according to the manufacturer's instructions with siRNA targeting DYRK3 (UAUAUGUGUAGAGCUUCUGGUACUC, Invitrogen). Experiments were carried out 72 hours post- transfection. shRNAs (shRNA-1 CCGGATGATCTTACAAACCTGCAAACTCGAGTTTGCAGGTTTGTAAGATC ATTTTTT, shRNA-2 CCGGGCCAGGGTCTATGATCACAAACTCGAGTTTGT-GATCATAGACCCTGGCTTTTT, shRNA-3 CCGGTGTTCAAATGTACTCTG-CAATCTCGAGATTGCAGAGTACATTTGAACATTTTT, shRNA-4 CCGGGTA-GGTCCAAATGCCAAGAAACTCGAGTTTCTTGGCATTTGGACCTACTTTTT)

were transfected using Lipofectamine 2000 according to the manufacturer's instructions.

Construction of plasmids, mutants and truncations

Full length DYRK3 cDNA clone (IRATp970D0168D, ImaGenes) was amplified using the primer 5´-CAC CAT GAA GTG GAA AGA GAA GTT GGG-3´ and 5´-CTA GCT AAT CAG TTT TGG CAA TAC AC-3´. The PCR product was inserted into the Gateway® system entry vector pENTR™/D-TOPO® (Invitrogen). The Gateway® LR recombination reaction was then performed between pENTR™/D-TOPO®-DYRK3 and its destination vector pcDNA5-FRT/TO-GFP or pcDNA5-FRT/TO-myc respectively (kindly provided by Manuel Bauer). The constructs pENTR-DYRK2, pENTR-DYRK3-K218M, pENTR-DYRK4, were a generous gift from by Markku Varjosalo (Institute of Molecular Systems Biology, ETH Zurich) and cloned into destination vectors as described above. Truncations of DYRK3 were created by amplification from the full length DYRK 3 using the primer pairs 5´-CAC CAT GAA GTG GAA AGA GAA GTT GGG-3´ and 5'-CTA TCG ATA AGC TAG ATG GTC TCG A-3' for DYRK3 (NT: 1-188), 5'-CAC CCG ATA TGA GGT GCT GAA A-3' and 5´-CTA GCT AAT CAG TTT TGG CAA TAC AC-3´ for DYRK3 (ΔNT: 189-568), 5'-CAC CAA ACG TGC CAA GTA CTT TAT T-3', subsequently insertion into pENTR™/D-TOPO® and cloning into designated destination vectors. pGEX-4T1-DYRK3 and pGEX-4T1-DYRK3-K218M for recombinant expression in *E.coli* were produced by amplifying full length wild type DYRK3 and DYRK3-K218M with the primer pair 5'-CCG GAA TTC ATG AAG TGG AAA GAG-3' and 5'- ATG CGG CCG CTC CTA GCT AAT CA-3' and insertion

into pGEX-4T/1-GST (Amersham) with BamHI and EcoRI. phrGFP-GW182 was kindly provided by Edward Chan (Department of Molecular and Experimental Medicine, The Scripps Research Institute, La Jolla). pEGFP-G3BP1 and pCI-puro-mRFP-DCP1a were from Paul J. Anderson (Division of Rheumatology, Immunology and Allergy, Brigham and Women's Hospital, Boston). pRK5-myc-PRAS40 was a gift from Do-Hyung Kim (Addgene #15476). pRK5-myc-RPTOR was a gift from David Sabatini (Addgene #1859). Cells were transiently transfected using Lipofectamine 2000 according to the manufacturer's instructions.

***In vitro* kinase assays**

Recombinant DYRK3 wild type and kinase-dead DYRK3-K218M GST-fusion proteins were produced in *E.coli* containing pGEX-4T-1-DYRK3 or pGEX-4T-1-DYRK3-K218M. Bacterial culture was harvested by centrifugation, lysed in PBS supplemented with 0.5 M NaCl, 2 mM EDTA, 10 mM DTT, 250 μM PMSF and EDTA free protease inhibitor (Roche) and sonicated. GST-fusion protein was purified using Glutathione sepharose 4B gravity-flow columns (GE Healthcare), eluted with 20 mM L-glutathione reduced in 50 mM Tris pH 8, 0.5 M NaCl and dialyzed against 50 mM Tris pH 8, 0.5 M NaCl. In vitro kinase assay was carried out in kinase buffer (25 mM Hepes, 5 mM $MgCl_2$, 5 mM $MnCl_2$, 0.5 mM DTT and 10 μM ATP), supplemented with 2 μCi γ^{33}P-ATP per reaction, 8 μg MBP and 2 μg DYRK3 for 30 minutes at room temperature. Proteins were separated by SDS-PAGE and γ^{33}P-incorporation was detected on a phosphorimager. In the in vitro kinase assay of DYRK3 and recombinant

PRAS40-myc, the phosphorylation of 0.5 μg PRAS40-myc was detected by separation of proteins with SDS-PAGE and immunoblotting as discribed.

Cell lysis and Immunoprecipitations

Cells were rinsed once with cold PBS and lysed in buffer (150 mM NaCl, 1% Triton X-100, 1 mM EDTA, 10 mM Tris pH 7.2, 10 μM PMSF, EDTA free protease inhibitor and PhosSTOP (Roche), 20 nM Calyculin A and 1 mM Na_3VO_4). For immunoprecipitations, cells were lysed in CHAPS buffer (120 mM NaCl, 0.3% CHAPS, 2 mM EDTA, 40 mM HEPES (pH 7.4), EDTA free protease inhibitor and PhosSTOP (Roche)) and cleared by centrifugation at 13,000 rpm for 10 minutes at 4°C. Primary antibody was added and incubated for 90 minutes with rotation. Immuno-complexes were captured by addition of protein G Sepharose slurry and incubation for 2 hours, followed by 7 wash steps with CHAPS buffer. Proteins were denatured by the addition sample buffer and boiling for 5 min. Proteins were resolved by SDS-PAGE and analyzed by immunoblotting as described.

FRAP analysis

HeLa cells were grown on 18mm cover slips, transfected with either pcDNA5-FRT/TO-GFP-DYRK3 or pcDNA5-FRT/TO-GFP-DYRK3-K218M using Lipofectamine 2000. Cells were monitored on a Leica SP5 Mid UV-VIS equipped with a 63x 1.4-0.6NA DIC, Oil, HCX Plan-Apochromat at a frame rate of 1 frame per 2.6 seconds. A defined region was bleached 3 times at full laser power using the 488 nm and 405 nm laser line after 5 frames. Recovery was monitored over a minimum of 150 seconds. FRAP analysis was carried out on using ImageJ (http://rsb.info.nih.gov/ij/). Recovery curves were fitted

and values were calculated with the formula $f(t) = f_0 + (f_{max} - f_0)*(1 - \exp(-\tau t))$ using Prism5 (GraphPad).

Relative quantification of transcripts

RNA was exctracted using Trizol reagent (Invitrogen) according to the manufacturer's instruction 72 hours post transfection. Reverse transcription was performed using the QuantiTect Reverse Transcription Kit (Qiagen). Quantitative real-time RT- PCR was carried out in an Mx3005P QPCR System (Stratagene) using primers against DYRK3 (5'-TCC TTC TGA ACC ACC TCC AC-3' and 5'-CCT TCA TCT CAC CTC CAT CC-3') and RPL13A (5'-GAG AAA GCC AAG ATC CAC TAC C-3' and 5'- TTG AGG ACC TCT GTG TAT TTG TC-3'). Relative change was calculated to control siRNA treatment and by using RPL13A as an internal standard for each sample.

Acknowledgements

We acknowledge all members of the Pelkmans lab for help and discussions during the course of this work and Lilli Stergiou and Herbert Polzhofer for help with experiments. F.W. is supported by a ProDoc fellowship from the Swiss National Science Foundation and is a member of the Life Science Zurich Graduate School. B.B. was supported by fellowships of the Boehringer Ingelheim Fonds and the Swiss National Science Foundation (SNF). L.P. is supported by the University of Zurich, the Swiss National Science Foundation and the European Union. This work was supported by the SystemsX.ch RTD

project PhosphoNetX.

References

Alvarez, M., Estivill, X., and de la Luna, S. (2003). DYRK1A accumulates in splicing speckles through a novel targeting signal and induces speckle disassembly. J Cell Sci *116*, 3099-3107.

Anderson, P., and Kedersha, N. (2002). Stressful initiations. J Cell Sci *115*, 3227-3234.

Anderson, P., and Kedersha, N. (2008). Stress granules: the Tao of RNA triage. Trends in biochemical sciences *33*, 141-150.

Anderson, P., and Kedersha, N. (2009). RNA granules: post-transcriptional and epigenetic modulators of gene expression. Nat Rev Mol Cell Biol *10*, 430-436.

Aranda, S., Laguna, A., and de la Luna, S. (2010). DYRK family of protein kinases: evolutionary relationships, biochemical properties, and functional roles. FASEB J.

Bain, J., Plater, L., Elliott, M., Shpiro, N., Hastie, C.J., McLauchlan, H., Klevernic, I., Arthur, J.S., Alessi, D.R., and Cohen, P. (2007). The selectivity of protein kinase inhibitors: a further update. Biochem J *408*, 297-315.

Baltz, A.G., Munschauer, M., Schwanhausser, B., Vasile, A., Murakawa, Y., Schueler, M., Youngs, N., Penfold-Brown, D., Drew, K., Milek, M.*, et al.* (2012). The mRNA-Bound Proteome and Its Global Occupancy Profile on Protein-Coding Transcripts. Mol Cell *46*, 674-690.

Bodenmiller, B., Mueller, L.N., Mueller, M., Domon, B., and Aebersold, R. (2007). Reproducible isolation of distinct, overlapping segments of the phosphoproteome. Nat Methods *4*, 231-237.

Bodenmiller, B., Wanka, S., Kraft, C., Urban, J., Campbell, D., Pedrioli, P.G., Gerrits, B., Picotti, P., Lam, H., Vitek, O.*, et al.* (2010). Phosphoproteomic analysis reveals interconnected system-wide responses to perturbations of kinases and phosphatases in yeast. Sci Signal *3*, rs4.

Boulay, A., Zumstein-Mecker, S., Stephan, C., Beuvink, I., Zilbermann, F., Haller, R., Tobler, S., Heusser, C., O'Reilly, T., Stolz, B.*, et al.* (2004). Antitumor efficacy of intermittent treatment schedules with the rapamycin derivative RAD001 correlates with prolonged inactivation of ribosomal protein S6 kinase 1 in peripheral blood mononuclear cells. Cancer Res *64*, 252-261.

Brangwynne, C.P., Eckmann, C.R., Courson, D.S., Rybarska, A., Hoege, C., Gharakhani, J., Julicher, F., and Hyman, A.A. (2009). Germline P granules are liquid droplets that localize by controlled dissolution/condensation. Science *324*, 1729-1732.

Brunn, G.J., Hudson, C.C., Sekulic, A., Williams, J.M., Hosoi, H., Houghton, P.J., Lawrence, J.C., Jr., and Abraham, R.T. (1997). Phosphorylation of the translational repressor PHAS-I by the mammalian target of rapamycin. Science *277*, 99-101.

Buchan, J.R., and Parker, R. (2009). Eukaryotic stress granules: the ins and outs of translation. Mol Cell *36*, 932-941.

Burnett, P.E., Barrow, R.K., Cohen, N.A., Snyder, S.H., and Sabatini, D.M. (1998). RAFT1 phosphorylation of the translational regulators p70 S6 kinase and 4E-BP1. Proc Natl Acad Sci U S A *95*, 1432-1437.

Carpenter, A.E., Jones, T.R., Lamprecht, M.R., Clarke, C., Kang, I.H., Friman, O., Guertin, D.A., Chang, J.H., Lindquist, R.A., Moffat, J.*, et al.* (2006). CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biol *7*, R100.

Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B.M., Strein, C., Davey, N.E., Humphreys, D.T., Preiss, T., Steinmetz, L.M.*, et al.* (2012). Insights into RNA Biology from an Atlas of Mammalian mRNA-Binding Proteins. Cell *149*, 1393-1406.

Elvira, G., Wasiak, S., Blandford, V., Tong, X.K., Serrano, A., Fan, X., del Rayo Sanchez-Carbente, M., Servant, F., Bell, A.W., Boismenu, D.*, et al.* (2006). Characterization of an RNA granule from developing brain. Molecular & cellular proteomics : MCP *5*, 635-651.

Erickson-Miller, C.L., Creasy, C., Chadderton, A., Hopson, C.B., Valoret, E.I., Gorczyca, M., Elefante, L., Wojchowski, D.M., Chomo, M., Fitch, D.M.*, et al.* (2007). GSK626616: A DYRK3 Inhibitor as a Potential New Therapy for the Treatment of Anemia. ASH Annual Meeting Abstracts *110*, 510.

Eulalio, A., Behm-Ansmant, I., and Izaurralde, E. (2007). P bodies: at the crossroads of post-transcriptional pathways. Nat Rev Mol Cell Biol *8*, 9-22.

Eystathioy, T., Jakymiw, A., Chan, E.K., Seraphin, B., Cougot, N., and Fritzler, M.J. (2003). The GW182 protein colocalizes with mRNA degradation associated proteins hDcp1 and hLSm4 in cytoplasmic GW bodies. RNA *9*, 1171-1173.

Gockler, N., Jofre, G., Papadopoulos, C., Soppa, U., Tejedor, F.J., and Becker, W. (2009). Harmine specifically inhibits protein kinase DYRK1A and interferes with neurite formation. FEBS J *276*, 6324-6337.

Hachet, O., Berthelot-Grosjean, M., Kokkoris, K., Vincenzetti, V., Moosbrugger, J., and Martin, S.G. (2011). A Phosphorylation Cycle Shapes Gradients of the DYRK Family Kinase Pom1 at the Plasma Membrane. Cell *145*, 1116-1128.

Hsu, P.P., Kang, S.A., Rameseder, J., Zhang, Y., Ottina, K.A., Lim, D., Peterson, T.R., Choi, Y., Gray, N.S., Yaffe, M.B.*, et al.* (2011). The mTOR-regulated phosphoproteome reveals a mechanism of mTORC1-mediated inhibition of growth factor signaling. Science *332*, 1317-1322.

Huber, A., Bodenmiller, B., Uotila, A., Stahl, M., Wanka, S., Gerrits, B., Aebersold, R., and Loewith, R. (2009). Characterization of the rapamycin-sensitive phosphoproteome reveals that Sch9 is a central coordinator of protein synthesis. Genes Dev *23*, 1929-1943.

Hyman, A.A., and Simons, K. (2012). Cell biology. Beyond oil and water--phase transitions in cells. Science *337*, 1047-1049.

Inoki, K., Li, Y., Xu, T., and Guan, K.L. (2003). Rheb GTPase is a direct target of TSC2 GAP activity and regulates mTOR signaling. Genes Dev *17*, 1829-1834.

Karaman, M.W., Herrgard, S., Treiber, D.K., Gallant, P., Atteridge, C.E., Campbell, B.T., Chan, K.W., Ciceri, P., Davis, M.I., Edeen, P.T.*, et al.* (2008). A quantitative analysis of kinase inhibitor selectivity. Nature biotechnology *26*, 127-132.

Kato, M., Han, T.W., Xie, S., Shi, K., Du, X., Wu, L.C., Mirzaei, H., Goldsmith, E.J., Longgood, J., Pei, J.*, et al.* (2012). Cell-free Formation of RNA Granules: Low Complexity Sequence Domains Form Dynamic Fibers within Hydrogels. Cell *149*, 753-767.

Kedersha, N.L., Gupta, M., Li, W., Miller, I., and Anderson, P. (1999). RNA-binding proteins TIA-1 and TIAR link the phosphorylation of eIF-2 alpha to the assembly of mammalian stress granules. J Cell Biol *147*, 1431-1442.

Loewith, R., and Hall, M.N. (2011). Target of rapamycin (TOR) in nutrient signaling and growth control. Genetics *189*, 1177-1201.

Ma, X.M., and Blenis, J. (2009). Molecular mechanisms of mTOR-mediated translational control. Nat Rev Mol Cell Biol *10*, 307-318.

Moriya, H., Shimizu-Yoshida, Y., Omori, A., Iwashita, S., Katoh, M., and Sakai, A. (2001). Yak1p, a DYRK family kinase, translocates to the nucleus and phosphorylates yeast Pop2p in response to a glucose signal. Genes Dev *15*, 1217-1228.

Ohn, T., Kedersha, N., Hickman, T., Tisdale, S., and Anderson, P. (2008). A functional RNAi screen links O-GlcNAc modification of ribosomal proteins to stress granule and processing body assembly. Nat Cell Biol *10*, 1224-1231.

Pang, K.M., Ishidate, T., Nakamura, K., Shirayama, M., Trzepacz, C., Schubert, C.M., Priess, J.R., and Mello, C.C. (2004). The minibrain kinase homolog, mbk-2, is required for spindle positioning and asymmetric cell division in early C. elegans embryos. Dev Biol *265*, 127-139.

Ramo, P., Sacher, R., Snijder, B., Begemann, B., and Pelkmans, L. (2009). CellClassifier: supervised learning of cellular phenotypes. Bioinformatics *25*, 3028-3030.

Sancak, Y., Thoreen, C.C., Peterson, T.R., Lindquist, R.A., Kang, S.A., Spooner, E., Carr, S.A., and Sabatini, D.M. (2007). PRAS40 is an insulin-regulated inhibitor of the mTORC1 protein kinase. Mol Cell *25*, 903-915.

Seifert, A., and Clarke, P.R. (2009). p38alpha- and DYRK1A-dependent phosphorylation of caspase-9 at an inhibitory site in response to hyperosmotic stress. Cellular signalling *21*, 1626-1633.

Sengupta, S., Peterson, T.R., and Sabatini, D.M. (2010). Regulation of the mTOR complex 1 pathway by nutrients, growth factors, and stress. Mol Cell *40*, 310-322.

Stitzel, M.L., Pellettieri, J., and Seydoux, G. (2006). The C. elegans DYRK Kinase MBK-2 Marks Oocyte Proteins for Degradation in Response to Meiotic Maturation. Current biology : CB *16*, 56-62.

Taira, N., Nihira, K., Yamaguchi, T., Miki, Y., and Yoshida, K. (2007). DYRK2 is targeted to the nucleus and controls p53 via Ser46 phosphorylation in the apoptotic response to DNA damage. Mol Cell *25*, 725-738.

Takahara, T., and Maeda, T. (2012). Transient Sequestration of TORC1 into Stress Granules during Heat Stress. Mol Cell.

Taminato, A., Bagattini, R., Gorjao, R., Chen, G., Kuspa, A., and Souza, G.M. (2002). Role for YakA, cAMP, and protein kinase A in regulation of stress responses of Dictyostelium discoideum cells. Mol Biol Cell *13*, 2266-2275.

Vander Haar, E., Lee, S.I., Bandhakavi, S., Griffin, T.J., and Kim, D.H. (2007). Insulin signalling to mTOR mediated by the Akt/PKB substrate PRAS40. Nat Cell Biol *9*, 316-323.

Weber, S.C., and Brangwynne, C.P. (2012). Getting RNA and protein in phase. Cell *149*, 1188-1191.

Yu, Y., Yoon, S.O., Poulogiannis, G., Yang, Q., Ma, X.M., Villen, J., Kubica, N., Hoffman, G.R., Cantley, L.C., Gygi, S.P.*, et al.* (2011). Phosphoproteomic analysis identifies Grb10 as an mTORC1 substrate that negatively regulates insulin signaling. Science *332*, 1322-1326.

Zhang, D., Li, K., Erickson-Miller, C.L., Weiss, M., and Wojchowski, D.M. (2005). DYRK gene structure and erythroid-restricted features of DYRK3 gene expression. Genomics *85*, 117-130.

Zoncu, R., Efeyan, A., and Sabatini, D.M. (2010). mTOR: from growth signal integration to cancer, diabetes and ageing. Nat Rev Mol Cell Biol *12*, 21-35.

Figure 1

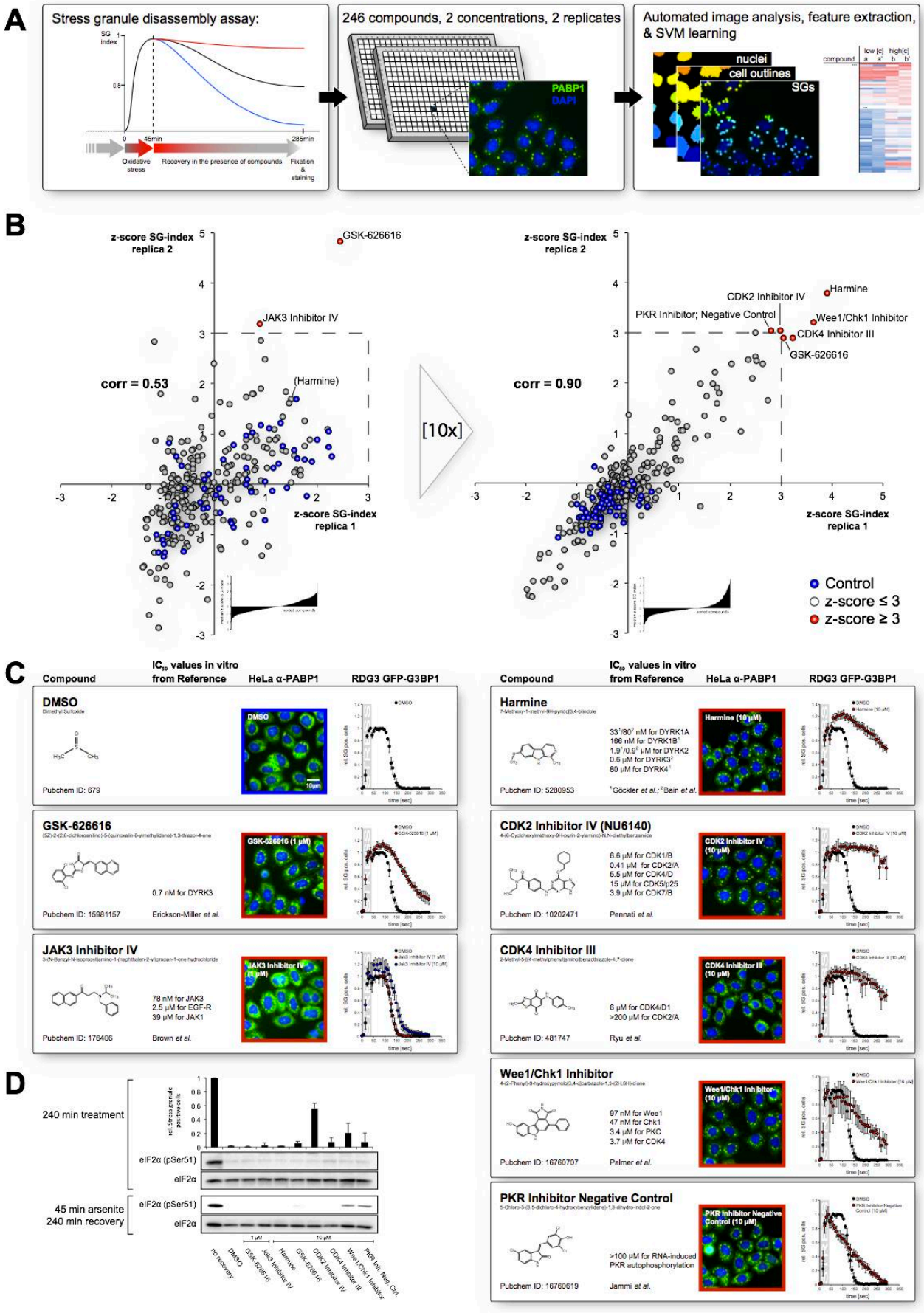

A
Stress granule disassembly assay:
SG index
0
45min
285min
Oxidative stress
Recovery in the presence of compounds
Fixation & staining
246 compounds, 2 concentrations, 2 replicates
PABP1
DAPI
Automated image analysis, feature extraction, & SVM learning
nuclei
cell outlines
SGs
compound
low [c] high [c]
a a' b b'
B
z-score SG-index replica 2
z-score SG-index replica 1
GSK-626616
JAK3 Inhibitor IV
(Harmine)
corr = 0.53
[10x]
corr = 0.90
CDK2 Inhibitor IV
Harmine
PKR Inhibitor; Negative Control
Wee1/Chk1 Inhibitor
CDK4 Inhibitor III
GSK-626616
Control
z-score ≤ 3
z-score ≥ 3
C
Compound
IC50 values in vitro from Reference
HeLa α-PABP1
RDG3 GFP-G3BP1
DMSO
Dimethyl Sulfoxide
Pubchem ID: 679
GSK-626616
0.7 nM for DYRK3
Pubchem ID: 15981157
Erickson-Miller et al.
JAK3 Inhibitor IV
78 nM for JAK3
2.5 µM for EGF-R
39 µM for JAK1
Pubchem ID: 176406
Brown et al.
Harmine
33[1]/80[2] nM for DYRK1A
166 nM for DYRK1B[1]
1.9[1]/0.9[2] µM for DYRK2
0.6 µM for DYRK3[2]
80 µM for DYRK4[1]
Pubchem ID: 5280953
[1]Göckler et al.; [2]Bain et al.
CDK2 Inhibitor IV (NU6140)
6.6 µM for CDK1/B
0.41 µM for CDK2/A
5.5 µM for CDK4/D
15 µM for CDK5/p25
3.9 µM for CDK7/B
Pubchem ID: 10202471
Pennati et al.
CDK4 Inhibitor III
6 µM for CDK4/D1
>200 µM for CDK2/A
Pubchem ID: 481747
Ryu et al.
Wee1/Chk1 Inhibitor
97 nM for Wee1
47 nM for Chk1
3.4 µM for PKC
3.7 µM for CDK4
Pubchem ID: 16760707
Palmer et al.
PKR Inhibitor Negative Control
>100 µM for RNA-induced PKR autophosphorylation
Pubchem ID: 16760619
Jammi et al.
D
240 min treatment
45 min arsenite
240 min recovery
rel. Stress granule positive cells
eIF2α (pSer51)
eIF2α
1 µM
10 µM
no recovery
DMSO
GSK-626616
JAK3 Inhibitor IV
Harmine
GSK-626616
CDK2 Inhibitor IV
CDK4 Inhibitor III
Wee1/Chk1 Inhibitor
PKR Inh. Neg. Ctrl.

Figure 2

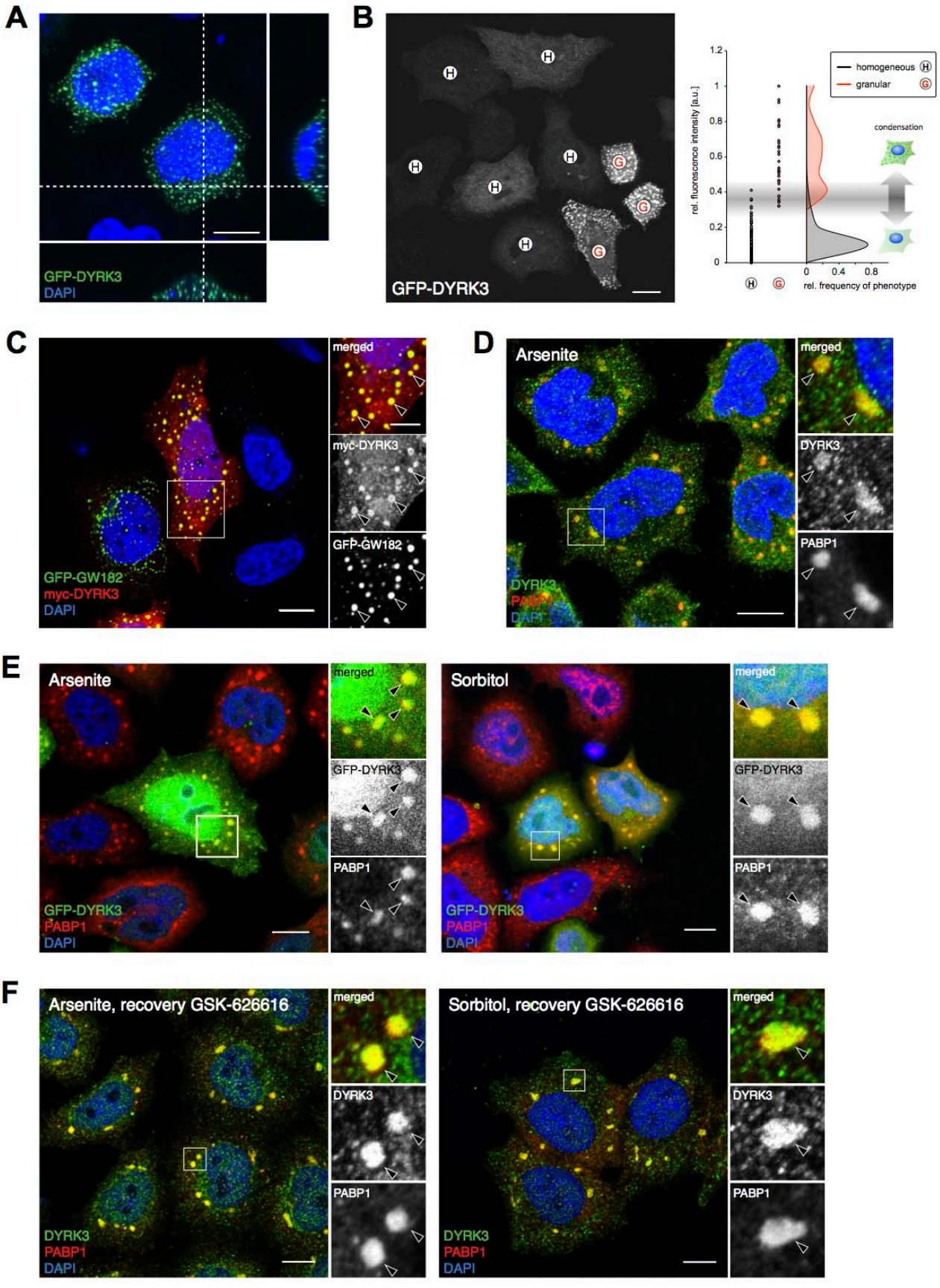

A
GFP-DYRK3
DAPI
B
GFP-DYRK3
homogeneous
granular
condensation
rel. fluorescence intensity [a.u.]
rel. frequency of phenotype
C
merged
myc-DYRK3
GFP-GW182
GFP-GW182
myc-DYRK3
DAPI
D
Arsenite
merged
DYRK3
PABP1
DYRK3
PABP1
DAPI
E
Arsenite
merged
GFP-DYRK3
PABP1
GFP-DYRK3
PABP1
DAPI
Sorbitol
merged
GFP-DYRK3
PABP1
GFP-DYRK3
PABP1
DAPI
F
Arsenite, recovery GSK-626616
merged
DYRK3
PABP1
DYRK3
PABP1
DAPI
Sorbitol, recovery GSK-626616
merged
DYRK3
PABP1
DYRK3
PABP1
DAPI

Figure 3

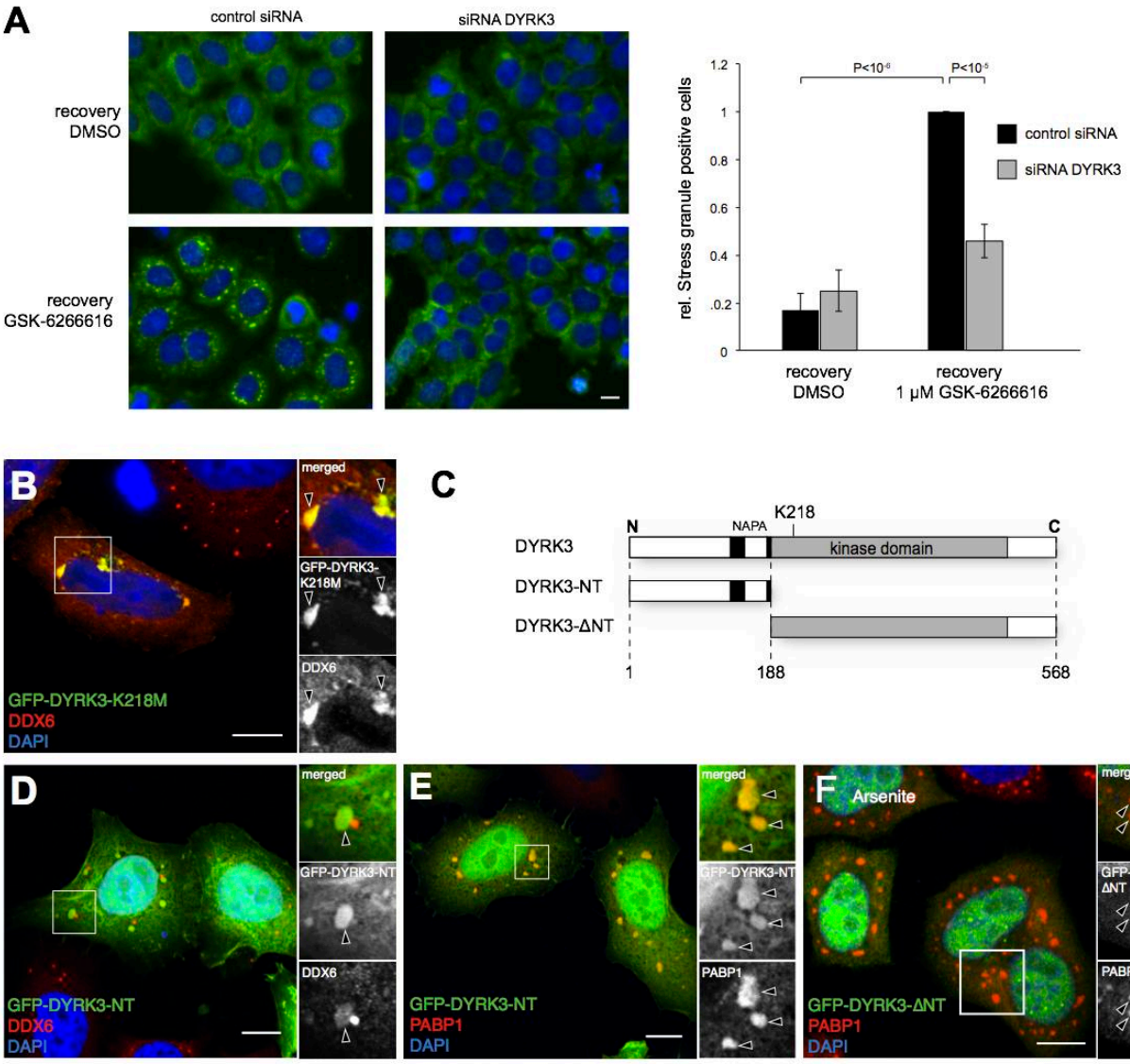
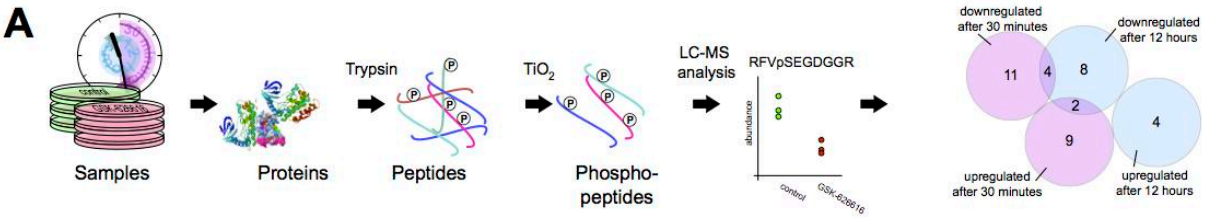
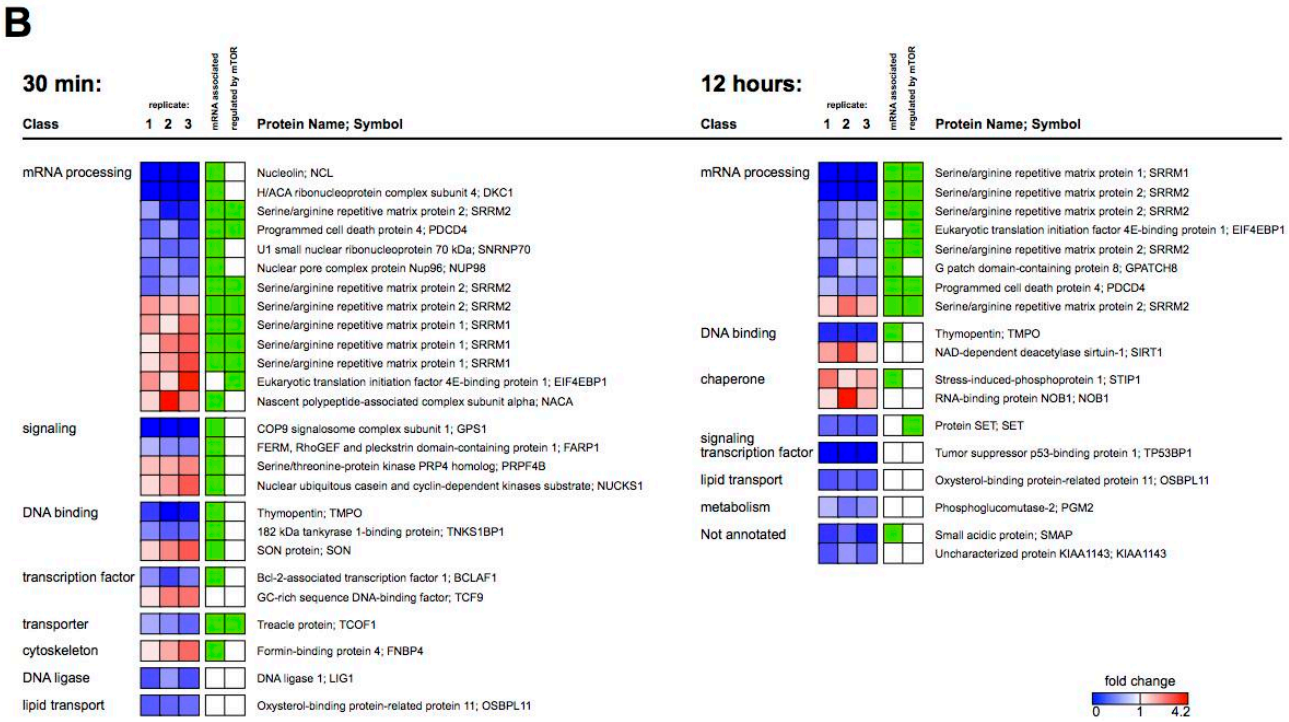
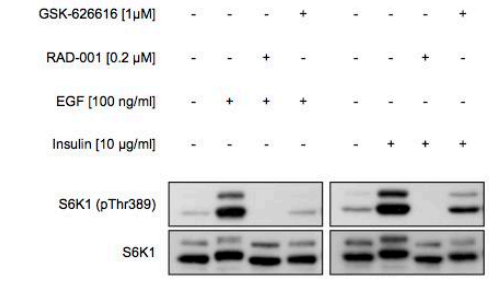

A
control siRNA
siRNA DYRK3
recovery DMSO
recovery GSK-6266616
rel. Stress granule positive cells
P<10-6
P<10-6
control siRNA
siRNA DYRK3
recovery DMSO
recovery 1 µM GSK-6266616
B
merged
GFP-DYRK3-K218M
DDX6
GFP-DYRK3-K218M
DDX6
DAPI
C
DYRK3
DYRK3-NT
DYRK3-ΔNT
N
NAPA
K218
kinase domain
C
1
188
568
D
merged
GFP-DYRK3-NT
DDX6
GFP-DYRK3-NT
DDX6
DAPI
E
merged
GFP-DYRK3-NT
PABP1
GFP-DYRK3-NT
PABP1
DAPI
F
Arsenite
merged
GFP-DYRK3-ΔNT
PABP1
GFP-DYRK3-ΔNT
PABP1
DAPI

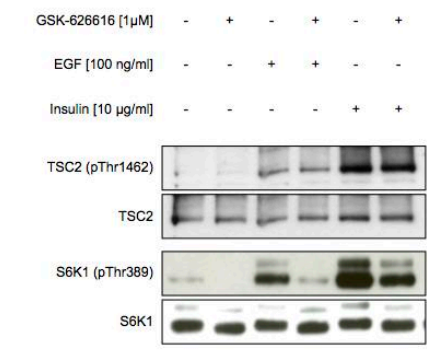
Figure 4

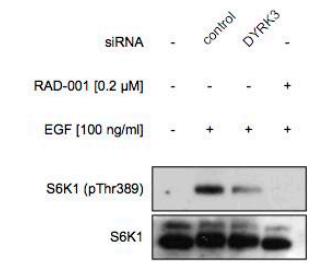

A
Samples
Proteins
Trypsin
Peptides
TiO2
Phospho-peptides
LC-MS analysis
RFVpSEGDGGR
abundance
control
GSK-626616
downregulated after 30 minutes
downregulated after 12 hours
upregulated after 30 minutes
upregulated after 12 hours
11
4
8
2
4
9

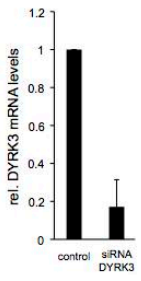

B
30 min:
12 hours:
Class
replicate: 1 2 3
mRNA associated
regulated by mTOR
Protein Name; Symbol
mRNA processing
Nucleolin; NCL
H/ACA ribonucleoprotein complex subunit 4; DKC1
Serine/arginine repetitive matrix protein 2; SRRM2
Programmed cell death protein 4; PDCD4
U1 small nuclear ribonucleoprotein 70 kDa; SNRNP70
Nuclear pore complex protein Nup96; NUP98
Serine/arginine repetitive matrix protein 2; SRRM2
Serine/arginine repetitive matrix protein 2; SRRM2
Serine/arginine repetitive matrix protein 1; SRRM1
Serine/arginine repetitive matrix protein 1; SRRM1
Serine/arginine repetitive matrix protein 1; SRRM1
Eukaryotic translation initiation factor 4E-binding protein 1; EIF4EBP1
Nascent polypeptide-associated complex subunit alpha; NACA
signaling
COP9 signalosome complex subunit 1; GPS1
FERM, RhoGEF and pleckstrin domain-containing protein 1; FARP1
Serine/threonine-protein kinase PRP4 homolog; PRPF4B
Nuclear ubiquitous casein and cyclin-dependent kinases substrate; NUCKS1
DNA binding
Thymopentin; TMPO
182 kDa tankyrase 1-binding protein; TNKS1BP1
SON protein; SON
transcription factor
Bcl-2-associated transcription factor 1; BCLAF1
GC-rich sequence DNA-binding factor; TCF9
transporter
Treacle protein; TCOF1
cytoskeleton
Formin-binding protein 4; FNBP4
DNA ligase
DNA ligase 1; LIG1
lipid transport
Oxysterol-binding protein-related protein 11; OSBPL11
mRNA processing
Serine/arginine repetitive matrix protein 1; SRRM1
Serine/arginine repetitive matrix protein 2; SRRM2
Serine/arginine repetitive matrix protein 2; SRRM2
Eukaryotic translation initiation factor 4E-binding protein 1; EIF4EBP1
Serine/arginine repetitive matrix protein 2; SRRM2
G patch domain-containing protein 8; GPATCH8
Programmed cell death protein 4; PDCD4
Serine/arginine repetitive matrix protein 2; SRRM2
DNA binding
Thymopentin; TMPO
NAD-dependent deacetylase sirtuin-1; SIRT1
chaperone
Stress-induced-phosphoprotein 1; STIP1
RNA-binding protein NOB1; NOB1
signaling transcription factor
Protein SET; SET
Tumor suppressor p53-binding protein 1; TP53BP1
lipid transport
Oxysterol-binding protein-related protein 11; OSBPL11
metabolism
Phosphoglucomutase-2; PGM2
Not annotated
Small acidic protein; SMAP
Uncharacterized protein KIAA1143; KIAA1143
fold change
0
1
4.2


C
GSK-626616 [1 µM]
RAD-001 [0.2 µM]
S6K1 (pThr389)
S6K1


D
GSK-626616 [1µM]
RAD-001 [0.2 µM]
EGF [100 ng/ml]
Insulin [10 µg/ml]
S6K1 (pThr389)
S6K1


E
GSK-626616 [1µM]
EGF [100 ng/ml]
Insulin [10 µg/ml]
TSC2 (pThr1462)
TSC2
S6K1 (pThr389)
S6K1


F
siRNA
control
DYRK3
RAD-001 [0.2 µM]
EGF [100 ng/ml]
S6K1 (pThr389)
S6K1


rel. DYRK3 mRNA levels
1.2
1
0.8
0.6
0.4
0.2
0
control
siRNA DYRK3

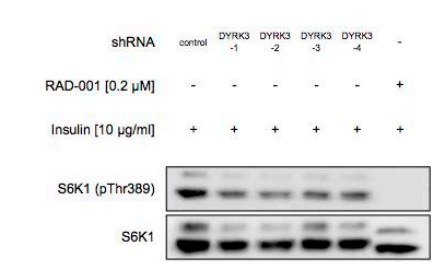
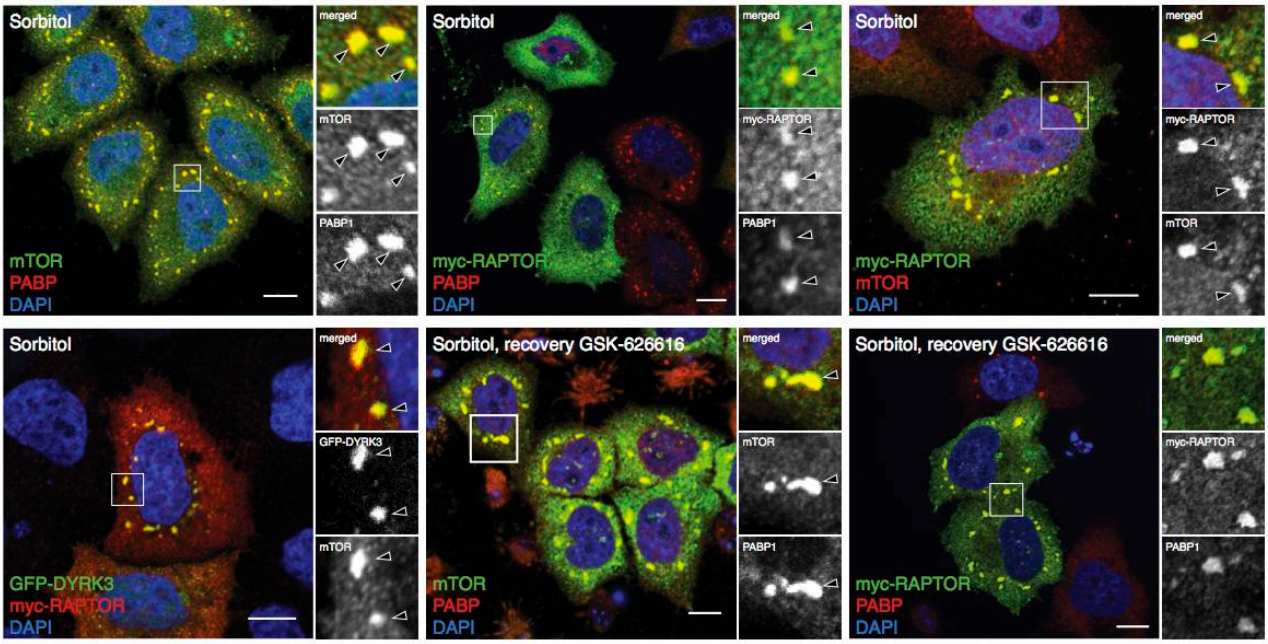
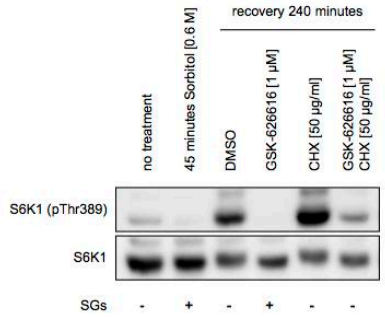

G
shRNA
control
DYRK3 -1
DYRK3 -2
DYRK3 -3
DYRK3 -4
RAD-001 [0.2 µM]
Insulin [10 µg/ml]
S6K1 (pThr389)
S6K1

Figure 5

A


Sorbitol
merged
mTOR
PABP1
mTOR
PABP
DAPI
Sorbitol
merged
myc-RAPTOR
PABP1
myc-RAPTOR
PABP
DAPI
Sorbitol
merged
myc-RAPTOR
mTOR
myc-RAPTOR
mTOR
DAPI
Sorbitol
merged
GFP-DYRK3
mTOR
GFP-DYRK3
myc-RAPTOR
DAPI
Sorbitol, recovery GSK-626616
merged
mTOR
PABP1
mTOR
PABP
DAPI
Sorbitol, recovery GSK-626616
merged
myc-RAPTOR
PABP1
myc-RAPTOR
PABP
DAPI

B


recovery 240 minutes
no treatment
45 minutes Sorbitol [0.6 M]
DMSO
GSK-626616 [1 µM]
CHX [50 µg/ml]
GSK-626616 [1 µM] CHX [50 µg/ml]
S6K1 (pThr389)
S6K1
SGs
- + - + - -

C

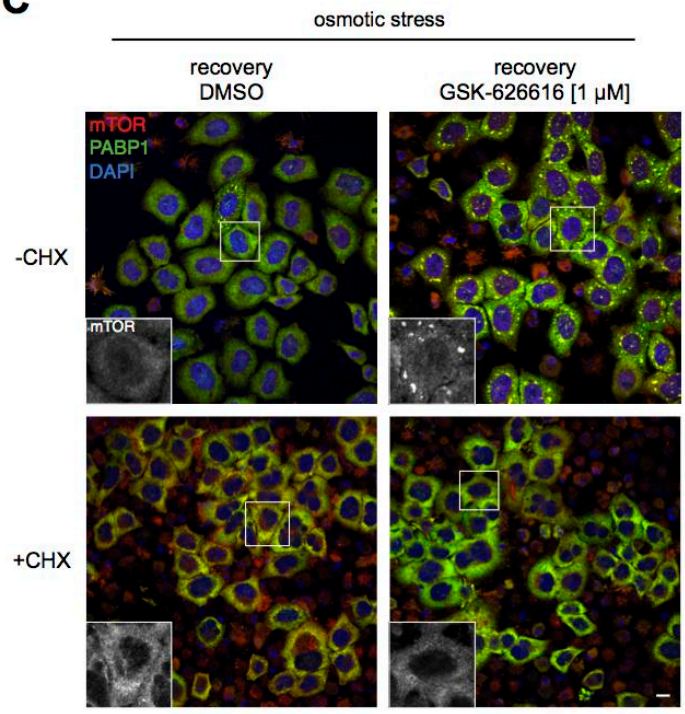
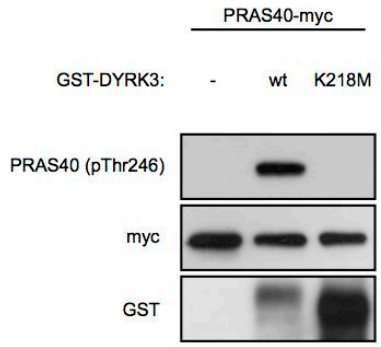
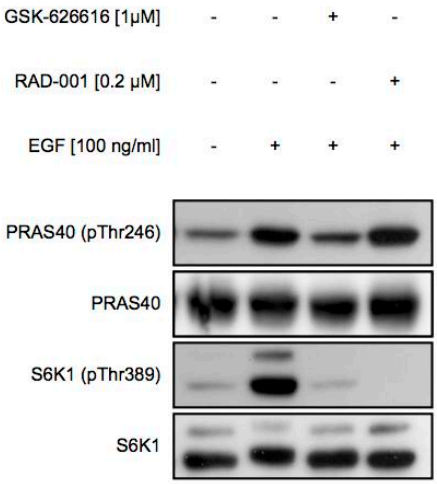

osmotic stress
recovery DMSO
recovery GSK-626616 [1 µM]
mTOR
PABP1
DAPI
-CHX
mTOR
+CHX

Figure 6

A


PRAS40-myc
GST-DYRK3: - wt K218M
PRAS40 (pThr246)
myc
GST

B


GSK-626616 [1μM] - - + -
RAD-001 [0.2 μM] - - - +
EGF [100 ng/ml] - + + +
PRAS40 (pThr246)
PRAS40
S6K1 (pThr389)
S6K1

C

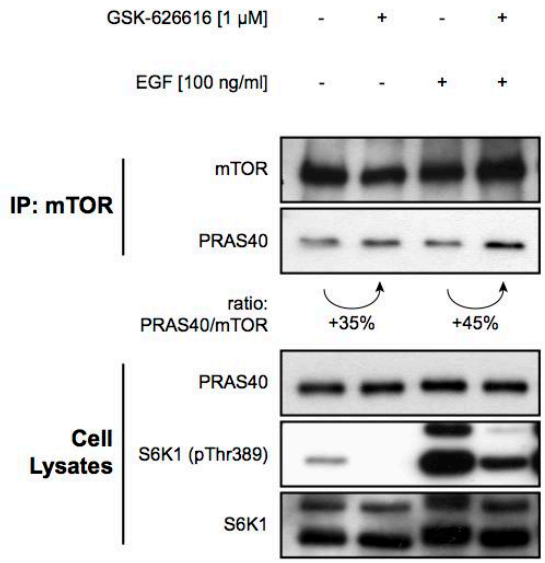
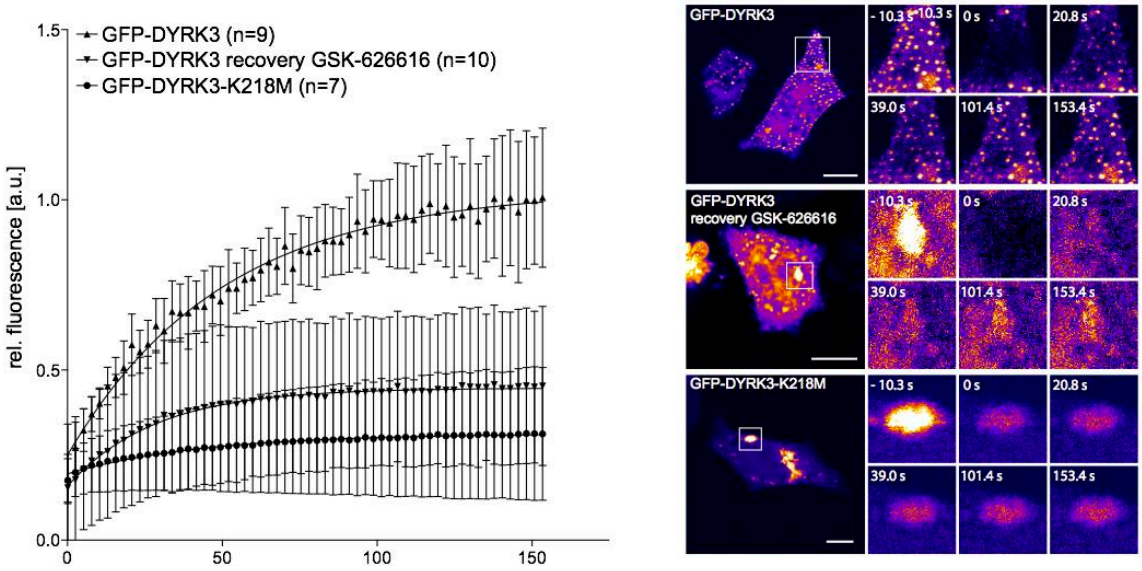

GSK-626616 [1 μM] - + - +
EGF [100 ng/ml] - - + +
IP: mTOR
mTOR
PRAS40
ratio:
PRAS40/mTOR
+35%
+45%
Cell
Lysates
PRAS40
S6K1 (pThr389)
S6K1

Figure 7

A


GFP-DYRK3 (n=9)
GFP-DYRK3 recovery GSK-626616 (n=10)
GFP-DYRK3-K218M (n=7)
rel. fluorescence [a.u.]
1.5
1.0
0.5
0.0
0
50
100
150
GFP-DYRK3
- 10.3 s
0 s
20.8 s
39.0 s
101.4 s
153.4 s
GFP-DYRK3
recovery GSK-626616
GFP-DYRK3-K218M

B

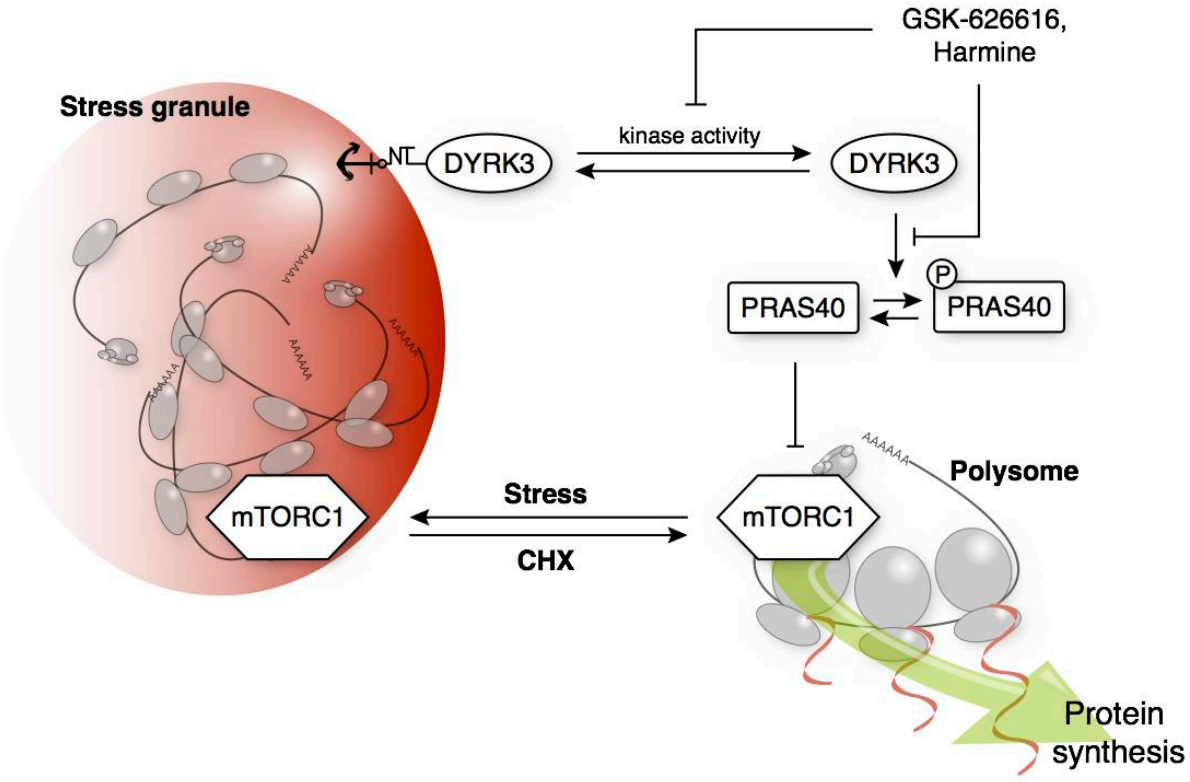
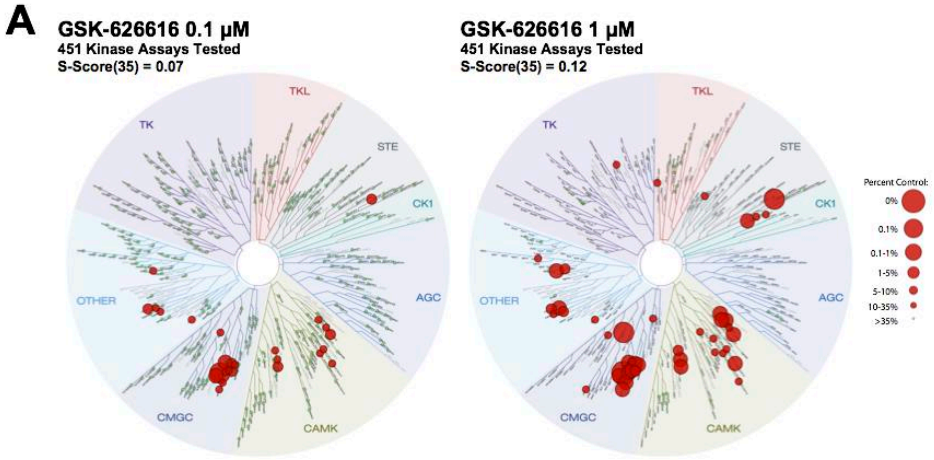
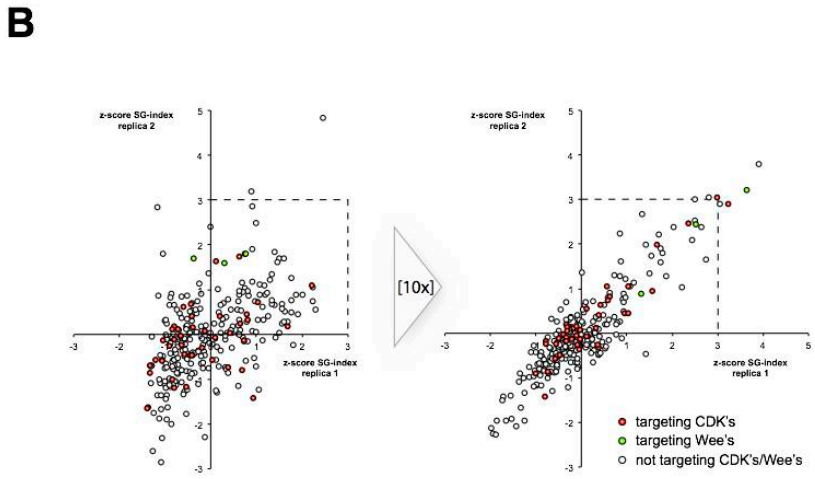
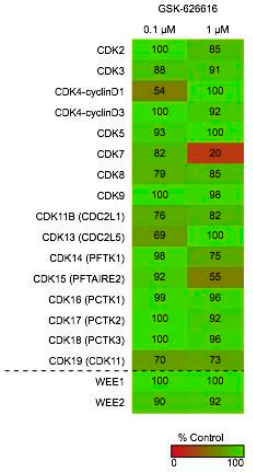

GSK-626616,
Harmine
Stress granule
NT
DYRK3
kinase activity
DYRK3
P
PRAS40
PRAS40
mTORC1
Stress
CHX
mTORC1
Polysome
Protein
synthesis

Figure S1


A
GSK-626616 0.1 µM
451 Kinase Assays Tested
S-Score(35) = 0.07
GSK-626616 1 µM
451 Kinase Assays Tested
S-Score(35) = 0.12
TKL
TK
STE
CK1
AGC
OTHER
CMGC
CAMK
Percent Control:
0%
0.1%
0.1-1%
1-5%
5-10%
10-35%
>35%


B
z-score SG-index replica 2
z-score SG-index replica 1
[10x]
targeting CDK's
targeting Wee's
not targeting CDK's/Wee's


C
GSK-626616
0.1 µM
1 µM
CDK2 100 85
CDK3 88 91
CDK4-cyclinD1 54 100
CDK4-cyclinD3 100 92
CDK5 93 100
CDK7 82 20
CDK8 79 85
CDK9 100 98
CDK11B (CDC2L1) 76 82
CDK13 (CDC2L5) 69 100
CDK14 (PFTK1) 98 75
CDK15 (PFTAIRE2) 92 55
CDK16 (PCTK1) 99 96
CDK17 (PCTK2) 100 92
CDK18 (PCTK3) 100 96
CDK19 (CDK11) 70 73
WEE1 100 100
WEE2 90 92
% Control
0
100

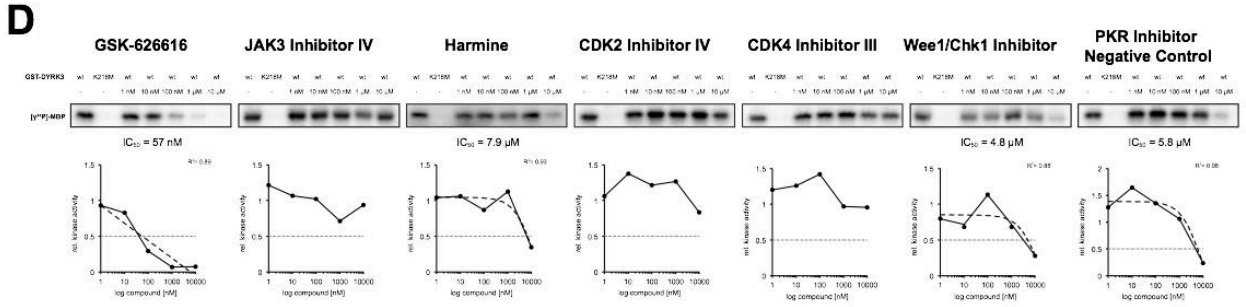

D
GSK-626616
JAK3 Inhibitor IV
Harmine
CDK2 Inhibitor IV
CDK4 Inhibitor III
Wee1/Chk1 Inhibitor
PKR Inhibitor Negative Control
GST-DYRK3
[γ-32P]-MBP
IC50 = 57 nM
IC50 = 7.9 µM
IC50 = 4.8 µM
IC50 = 5.8 µM
rel. kinase activity
log compound [nM]

Figure S2

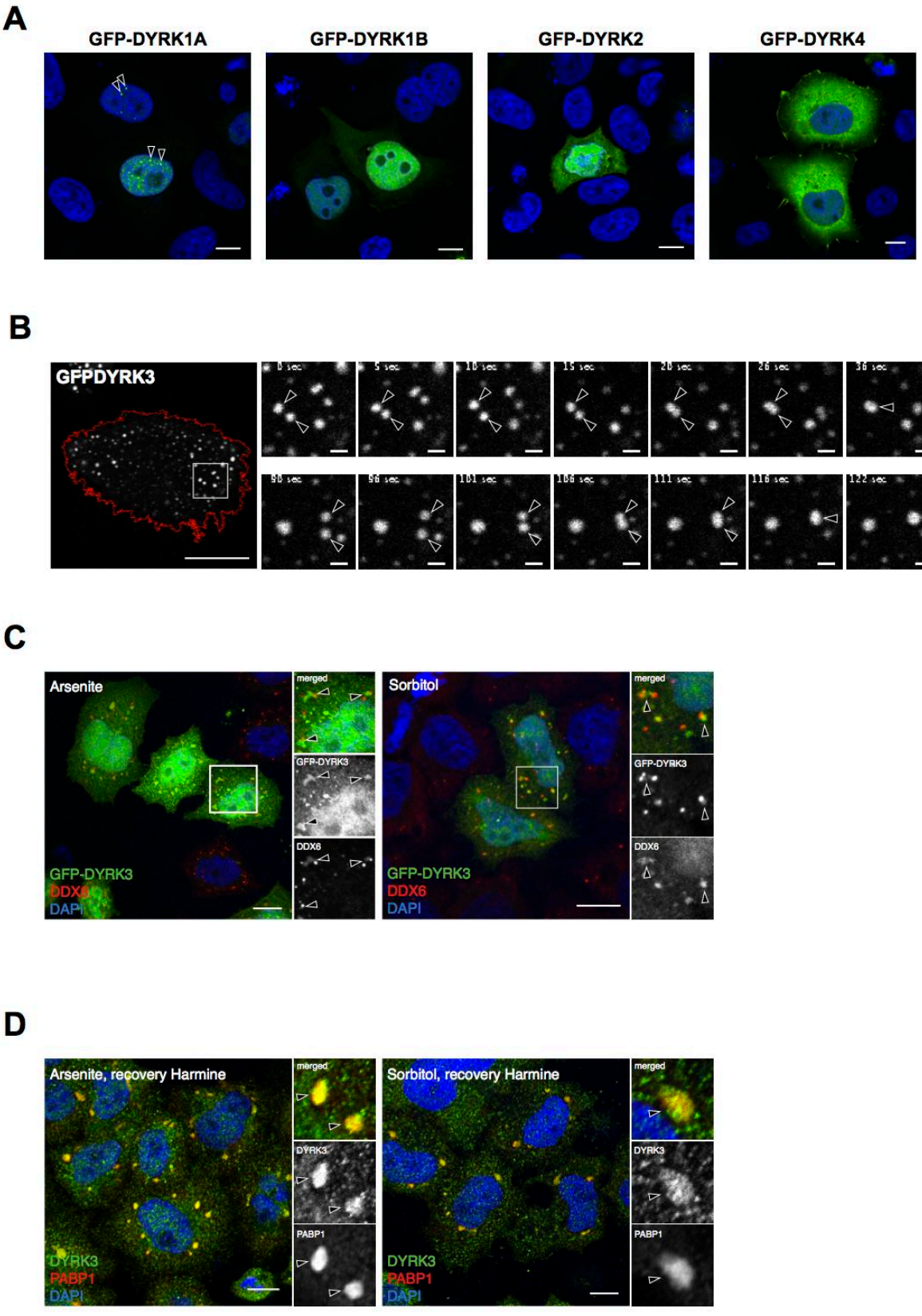
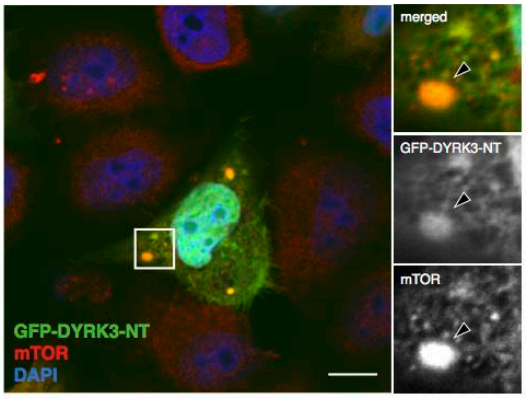
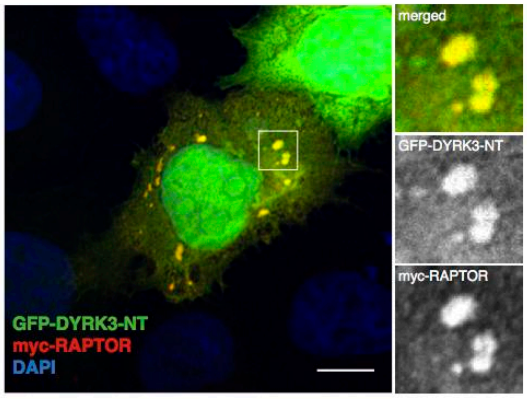

A
GFP-DYRK1A
GFP-DYRK1B
GFP-DYRK2
GFP-DYRK4
B
GFPDYRK3
C
Arsenite
merged
GFP-DYRK3
DDX6
GFP-DYRK3
DDX6
DAPI
Sorbitol
merged
GFP-DYRK3
DDX6
GFP-DYRK3
DDX6
DAPI
D
Arsenite, recovery Harmine
merged
DYRK3
PABP1
DYRK3
PABP1
DAPI
Sorbitol, recovery Harmine
merged
DYRK3
PABP1
DYRK3
PABP1
DAPI

Figure S3

A


merged
GFP-DYRK3-NT
mTOR
GFP-DYRK3-NT
mTOR
DAPI


merged
GFP-DYRK3-NT
myc-RAPTOR
GFP-DYRK3-NT
myc-RAPTOR
DAPI

B

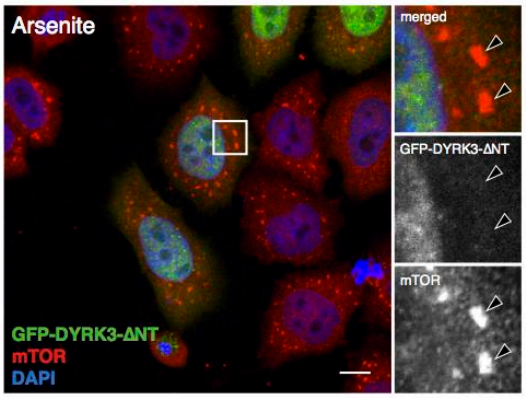

Arsenite
merged
GFP-DYRK3-ΔNT
mTOR
GFP-DYRK3-ΔNT
mTOR
DAPI

Figure S4

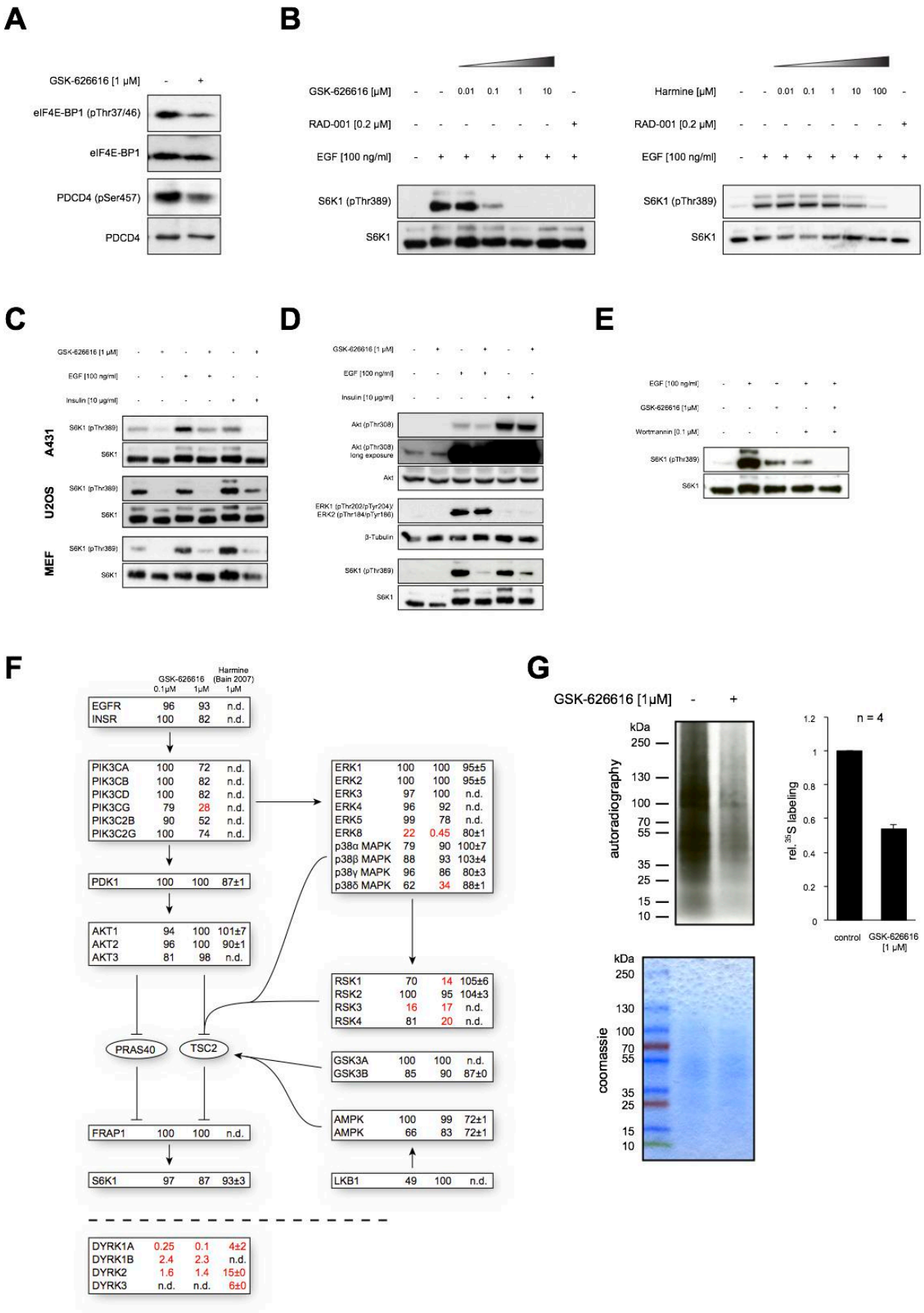
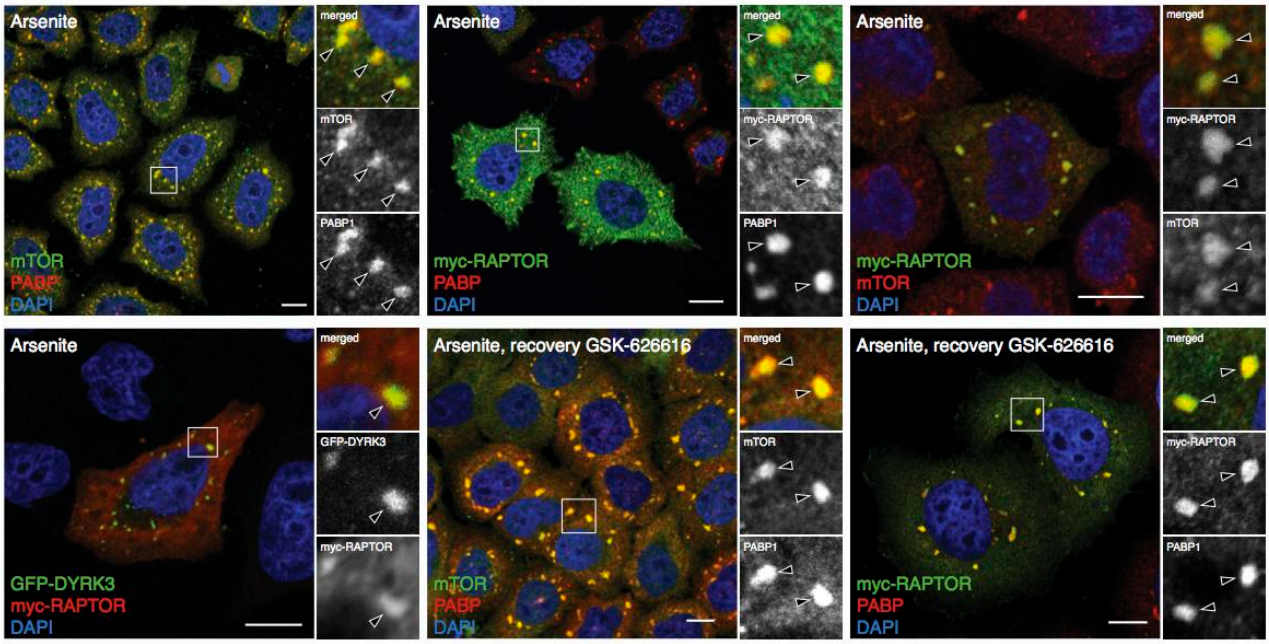
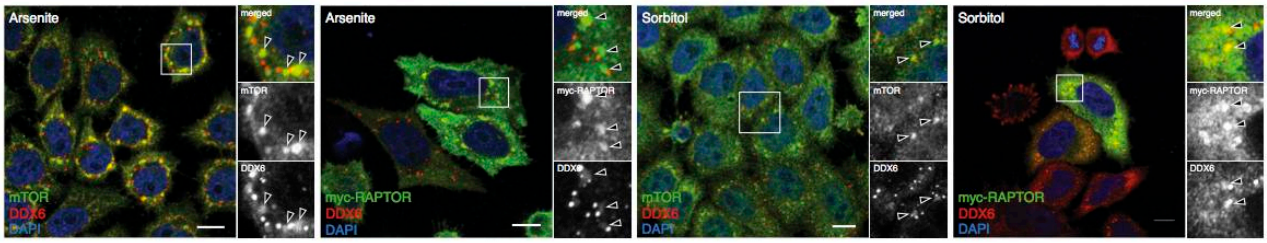

A
GSK-626616 [1 µM]
eIF4E-BP1 (pThr37/46)
eIF4E-BP1
PDCD4 (pSer457)
PDCD4
B
GSK-626616 [µM]
RAD-001 [0.2 µM]
EGF [100 ng/ml]
S6K1 (pThr389)
S6K1
Harmine [µM]
C
A431
U2OS
MEF
D
Akt (pThr308)
Akt (pThr308) long exposure
Akt
ERK1 (pThr202/pTyr204)/ ERK2 (pThr184/pTyr186)
β-Tubulin
E
Wortmannin [0.1 µM]
F
GSK-626616 0.1µM
GSK-626616 1µM
Harmine (Bain 2007) 1µM
EGFR 96 93 n.d.
INSR 100 82 n.d.
PIK3CA 100 72 n.d.
PIK3CB 100 82 n.d.
PIK3CD 100 82 n.d.
PIK3CG 79 28 n.d.
PIK3C2B 90 52 n.d.
PIK3C2G 100 74 n.d.
PDK1 100 100 87±1
AKT1 94 100 101±7
AKT2 96 100 90±1
AKT3 81 98 n.d.
PRAS40
TSC2
FRAP1 100 100 n.d.
S6K1 97 87 93±3
ERK1 100 100 95±5
ERK2 100 100 95±5
ERK3 97 100 n.d.
ERK4 96 92 n.d.
ERK5 99 78 n.d.
ERK8 22 0.45 80±1
p38α MAPK 79 90 100±7
p38β MAPK 88 93 103±4
p38γ MAPK 96 86 80±3
p38δ MAPK 62 34 88±1
RSK1 70 14 105±6
RSK2 100 95 104±3
RSK3 16 17 n.d.
RSK4 81 20 n.d.
GSK3A 100 100 n.d.
GSK3B 85 90 87±0
AMPK 100 99 72±1
AMPK 66 83 72±1
LKB1 49 100 n.d.
DYRK1A 0.25 0.1 4±2
DYRK1B 2.4 2.3 n.d.
DYRK2 1.6 1.4 15±0
DYRK3 n.d. n.d. 6±0
G
GSK-626616 [1µM]
kDa
autoradiography
coomassie
rel. 35S labeling
n = 4
control
GSK-626616 [1 µM]

Figure S5

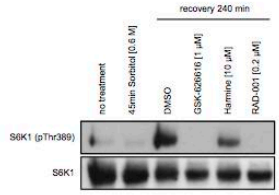
A



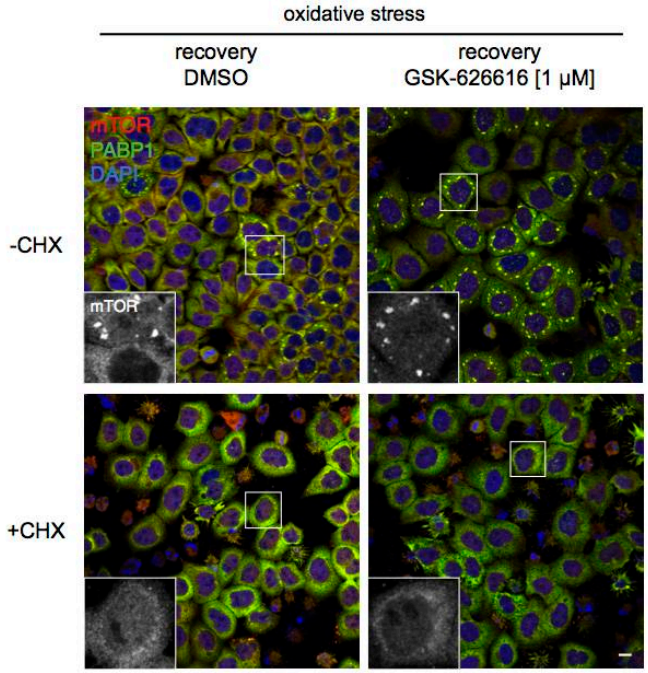
B



C



D



E

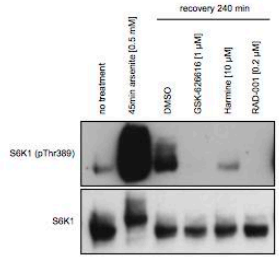
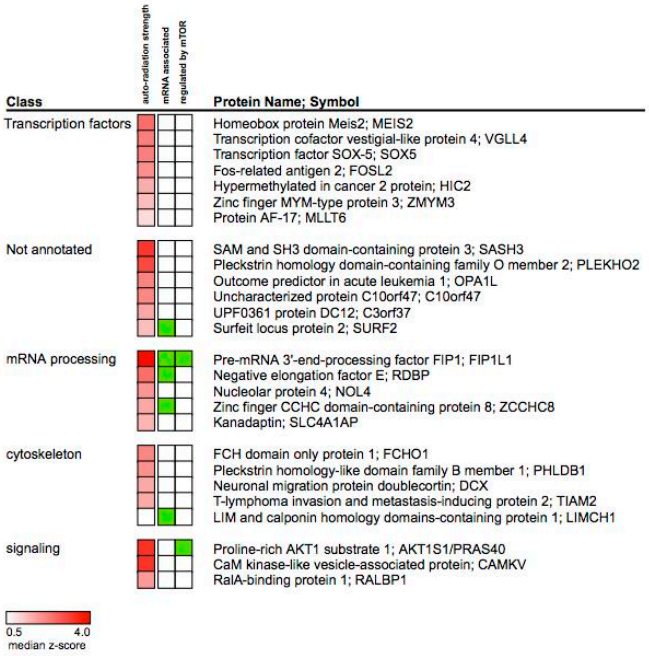
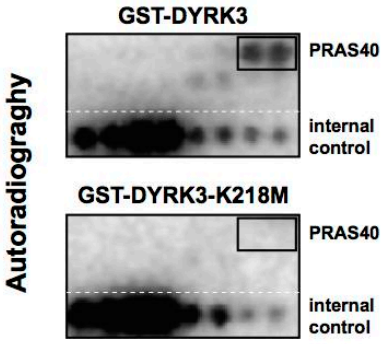


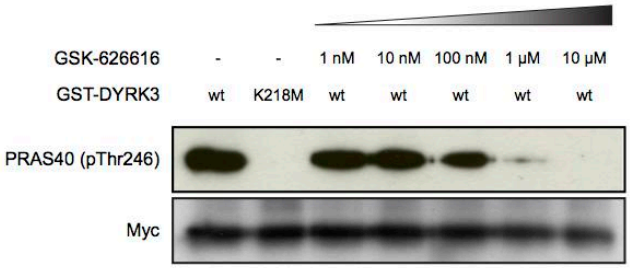
Figure S6

A


auto-radiation strength
mRNA associated
regulated by mTOR
Class
Protein Name; Symbol
Transcription factors
Homeobox protein Meis2; MEIS2
Transcription cofactor vestigial-like protein 4; VGLL4
Transcription factor SOX-5; SOX5
Fos-related antigen 2; FOSL2
Hypermethylated in cancer 2 protein; HIC2
Zinc finger MYM-type protein 3; ZMYM3
Protein AF-17; MLLT6
Not annotated
SAM and SH3 domain-containing protein 3; SASH3
Pleckstrin homology domain-containing family O member 2; PLEKHO2
Outcome predictor in acute leukemia 1; OPA1L
Uncharacterized protein C10orf47; C10orf47
UPF0361 protein DC12; C3orf37
Surfeit locus protein 2; SURF2
mRNA processing
Pre-mRNA 3'-end-processing factor FIP1; FIP1L1
Negative elongation factor E; RDBP
Nucleolar protein 4; NOL4
Zinc finger CCHC domain-containing protein 8; ZCCHC8
Kanadaptin; SLC4A1AP
cytoskeleton
FCH domain only protein 1; FCHO1
Pleckstrin homology-like domain family B member 1; PHLDB1
Neuronal migration protein doublecortin; DCX
T-lymphoma invasion and metastasis-inducing protein 2; TIAM2
LIM and calponin homology domains-containing protein 1; LIMCH1
signaling
Proline-rich AKT1 substrate 1; AKT1S1/PRAS40
CaM kinase-like vesicle-associated protein; CAMKV
RalA-binding protein 1; RALBP1
0.5
4.0
median z-score


Autoradiograghy
GST-DYRK3
PRAS40
internal control
GST-DYRK3-K218M
PRAS40
internal control

B


GSK-626616
- - 1 nM 10 nM 100 nM 1 µM 10 µM
GST-DYRK3
wt K218M wt wt wt wt wt
PRAS40 (pThr246)
Myc

BIBLIOGRAPHY

[1] Kritikou, E., Pulverer, B. & Heinrichs, A. All systems go! *Nature Reviews Molecular Cell Biology* **7**, 801–801 (2006).

[2] Wilson, I. Top-down versus bottom-up—rediscovering physiology via systems biology? *Molecular Systems Biology* **3** (2007).

[3] Tachibana, C. Systems Biology and Bioinformatics: Something for Everyone. *Science* (2010).

[4] Chuang, H., Hofree, M. & Ideker, T. A Decade of Systems Biology. *Annual Review of Cell and Developmental Biology* **26**, 721–744 (2010).

[5] O'Malley, M. & Dupré, J. Fundamental issues in systems biology. *BioEssays* **27**, 1270–1276 (2005).

[6] Bordel, S. & Nookaew, I. Systems Biology or just Biology? A report of the 11th International Conference on Systems Biology. *Biotechnology Journal* **5**, 1257–1260 (2010).

[7] Aebersold, R. Molecular Systems Biology: a new journal for a new biology? *Molecular Systems Biology* **1**, E1–E2 (2005).

[8] The nobel prize in physiology or medicine 1992 (1992). URL `http://www.nobelprize.org/nobel_prizes/medicine/laureates/1992/presentation-speech.html#`.

[9] Blum, H. Nobelpreis für Medizin 1992. *Dtsch med Wochenschr* **117**, 1935–1938 (1992).

[10] Krebs, E. & Fischer, E. Phosphorylase and related enzymes of glycogen metabolism. *Vitamins & Hormones* **22**, 399–410 (1964).

[11] Marx, J., Appenzeller, T., Amato, I. & Moffat, A. US researchers gather a bumper crop of laurels. *Science* **258**, 542–545 (1992).

[12] Manning, G., Whyte, D., Martinez, R., Hunter, T. & Sudarsanam, S. The protein kinase complement of the human genome. *Science* **298**, 1912–1934 (2002).

[13] Cohen, P. The regulation of protein function by multisite phosphorylation – a 25 year update. *Trends in Biochemical Sciences* **25**, 596–601 (2000).

[14] Lemeer, S. & Heck, A. The phosphoproteomics data explosion. *Current Opinion in Chemical Biology* **13**, 414–420 (2009).

[15] Stark, C. *et al.* PhosphoGRID: a database of experimentally verified in vivo protein phosphorylation sites from the budding yeast Saccharomyces cerevisiae. *Database: the Journal of Biological Databases and Duration* **2010** (2010).

[16] Burnett, G. & Kennedy, E. The enzymatic phosphorylation of proteins. *Journal of Biological Chemistry* **211**, 969–980 (1954).

[17] Alberts, B. *et al. Molecular biology of the cell, 4th edition* (Garland Science, New York, 2002).

[18] Griffiths, A. *Introduction to genetic analysis, 7th edition* (WH Freeman, 2000).

[19] Stem cell related products. URL `http://www.biocat.com/cgi-bin/page/sub1.pl?main_group=cell_biology&sub1=stem_cell_related_products`.

[20] The cell. URL `http://web.jjay.cuny.edu/~acarpi/NSC/13-cells.htm`.

[21] Pinna, L. & Ruzzene, M. How do protein kinases recognize their substrates? *Biochimica et biophysica acta* **1314**, 191–225 (1996).

[22] Lander, E. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).

[23] Dinkel, H. *et al.* Phospho.ELM: a database of phosphorylation sites–update 2011. *Nucleic Acids Research* **39**, D261–D267 (2011).

[24] Bodenmiller, B. *et al.* PhosphoPep–a database of protein phosphorylation sites in model organisms. *Nature Biotechnology* **26**, 1339–1340 (2008).

[25] Salazar, C. & Höfer, T. Multisite protein phosphorylation–from molecular mechanisms to kinetic models. *Febs Journal* **276**, 3177–3198 (2009).

[26] Rubin, G., Yandell, M., Wortman, J., Gabor, G. *et al.* Comparative genomics of the eukaryotes. *Science* **287**, 2204–2215 (2000).

[27] Venter, J. *et al.* The sequence of the human genome. *Science* **291**, 1304–1351 (2001).

[28] Hunter, T. & Plowman, G. The protein kinases of budding yeast: six score and more. *Trends in Biochemical Sciences* **22**, 18–22 (1997).

[29] Manning, G., Plowman, G., Hunter, T. & Sudarsanam, S. Evolution of protein kinase signaling from yeast to man. *Trends in Biochemical Sciences* **27**, 514–520 (2002).

[30] Morrison, D., Murakami, M. & Cleghon, V. Protein kinases and phosphatases in the drosophila genome. *The Journal of Cell Biology* **150**, F57–62 (2000).

[31] Zhu, H. *et al.* Analysis of yeast protein kinases using protein chips. *Nature Genetics* **26**, 283–290 (2000).

[32] Miranda-Saavedra, D. & Barton, G. Classification and functional annotation of eukaryotic protein kinases. *Proteins: Structure, Function, and Bioinformatics* **68**, 893–914 (2007).

[33] Cohen, P. The role of protein phosphorylation in human health and disease. *European Journal of Biochemistry* **268**, 5001–5010 (2001).

[34] Cohen, P. Protein kinases–the major drug targets of the twenty-first century? *Nature Reviews Drug Discovery* **1**, 309–315 (2002).

[35] Anderson, N. & Anderson, N. Proteome and proteomics: new technologies, new concepts, and new words. *Electrophoresis* **19**, 1853–1861 (1998).

[36] Mumby, M. & Brekken, D. Phosphoproteomics: new insights into cellular signaling. *Genome Biology* **6**, 230 (2005).

[37] Johnson, S. & Hunter, T. Phosphoproteomics finds its timing. *Nature Biotechnology* **22**, 1093–1094 (2004).

[38] Yates, J., Ruse, C. & Nakorchevsky, A. Proteomics by mass spectrometry: approaches, advances, and applications. *Annual Review of Biomedical Engineering* **11**, 49–79 (2009).

[39] Gstaiger, M. & Aebersold, R. Applying mass spectrometry-based proteomics to genetics, genomics and network biology. *Nature Reviews Genetics* **10**, 617–627 (2009).

[40] Bodenmiller, B., Mueller, L. N., Mueller, M., Domon, B. & Aebersold, R. Reproducible isolation of distinct, overlapping segments of the phosphoproteome. *Nature Methods* **4**, 231–237 (2007).

[41] Choudhary, C. & Mann, M. Decoding signalling networks by mass spectrometry–based proteomics. *Nature Reviews Molecular Cell Biology* **11**, 427–439 (2010).

[42] Beausoleil, S. *et al.* Large-scale characterization of HeLa cell nuclear phosphoproteins. *Proceedings of the National Academy of Sciences* **101**, 12130–12135 (2004).

[43] Villén, J., Beausoleil, S., Gerber, S. & Gygi, S. Large-scale phosphorylation analysis of mouse liver. *Proceedings of the National Academy of Sciences* **104**, 1488–1493 (2007).

[44] Zhai, B., Villén, J., Beausoleil, S., Mintseris, J. & Gygi, S. Phosphoproteome analysis of drosophila melanogaster embryos. *Journal of Proteome Research* **7**, 1675–1682 (2008).

[45] Wilson-Grady, J., Villén, J. & Gygi, S. Phosphoproteome analysis of fission yeast. *Journal of Proteome Research* **7**, 1088–1097 (2008).

[46] Swaney, D., Wenger, C., Thomson, J. & Coon, J. Human embryonic stem cell phosphoproteome revealed by electron transfer dissociation tandem mass spectrometry. *Proceedings of the National Academy of Sciences* **106**, 995–1000 (2009).

[47] Gnad, F. *et al.* PHOSIDA (phosphorylation site database): management, structural and evolutionary investigation, and prediction of phosphosites. *Genome Biology* **8**, R250 (2007).

[48] Hornbeck, P., Chabra, I., Kornhauser, J., Skrzypek, E. & Zhang, B. PhosphoSite: A bioinformatics resource dedicated to physiological protein phosphorylation. *Proteomics* **4**, 1551–1561 (2004).

[49] Gruhler, A. *et al.* Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. *Molecular & Cellular Proteomics* **4**, 310–327 (2005).

[50] Tao, W. *et al.* Quantitative phosphoproteome analysis using a dendrimer conjugation chemistry and tandem mass spectrometry. *Nature Methods* **2**, 591–598 (2005).

[51] Malik, R. *et al.* Quantitative analysis of the human spindle phosphoproteome at distinct mitotic stages. *Journal of Proteome Research* **8**, 4553–4563 (2009).

[52] Olsen, J. *et al.* Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis. *Science Signaling* **3**, ra3 (2010).

[53] Bantscheff, M., Schirle, M., Sweetman, G., Rick, J. & Kuster, B. Quantitative mass spectrometry in proteomics: a critical review. *Analytical and Bioanalytical Chemistry* **389**, 1017–1031 (2007).

[54] Nita-Lazar, A., Saito-Benz, H. & White, F. Quantitative phosphoproteomics by mass spectrometry: past, present, and future. *Proteomics* **8**, 4433–4443 (2008).

[55] Zhu, W., Smith, J. & Huang, C. Mass spectrometry-based label-free quantitative proteomics. *Journal of Biomedicine and Biotechnology* **2010** (2009).

[56] Mueller, L., Brusniak, M., Mani, D. & Aebersold, R. An assessment of software solutions for the analysis of mass spectrometry based

quantitative proteomics data. *Journal of Proteome Research* **7**, 51–61 (2008).

[57] Yang, C., He, Z. & Yu, W. Comparison of public peak detection algorithms for maldi mass spectrometry data analysis. *BMC Bioinformatics* **10**, 4 (2009).

[58] America, A. & Cordewener, J. Comparative LC-MS: A landscape of peaks and valleys. *Proteomics* **8**, 731–749 (2008).

[59] Mueller, L. N. *et al.* SuperHirn – a novel tool for high resolution LC-MS-based peptide/protein profiling. *Proteomics* **7**, 3470–3480 (2007).

[60] Sturm, M. *et al.* OpenMS–an open-source software framework for mass spectrometry. *BMC Bioinformatics* **9**, 163 (2008).

[61] May, D., Law, W., Fitzgibbon, M., Fang, Q. & McIntosh, M. Software platform for rapidly creating computational tools for mass spectrometry-based proteomics. *Journal of Proteome Research* **8**, 3212–3217 (2009).

[62] Li, X., Eugene, C., Kemp, C., Zhang, H. & Aebersold, R. A software suite for the generation and comparison of peptide arrays from sets of data collected by liquid chromatography-mass spectrometry. *Molecular & Cellular Proteomics* **4**, 1328–1340 (2005).

[63] Pluskal, T., Castillo, S., Villar-Briones, A. & Orešič, M. MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* **11**, 395 (2010).

[64] Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized ppb-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology* **26**, 1367–1372 (2008).

[65] MaxQuant for label-free quantification. URL `http://maxquant.org/`.

[66] Ubersax, J. *et al.* Targets of the cyclin-dependent kinase Cdk1. *Nature* **425**, 859–864 (2003).

[67] Dephoure, N., Howson, R., Blethrow, J., Shokat, K. & O'Shea, E. Combining chemical genetics and proteomics to identify protein kinase substrates. *Proceedings of the National Academy of Sciences* **102**, 17940–17945 (2005).

[68] Ptacek, J. *et al.* Global analysis of protein phosphorylation in yeast. *Nature* **438**, 679–684 (2005).

[69] Fiedler, D. *et al.* Functional organization of the S.cerevisiae phosphorylation network. *Cell* **136**, 952–963 (2009).

[70] Breitkreutz, A. *et al.* A global protein kinase and phosphatase interaction network in yeast. *Science* **328**, 1043–1046 (2010).

[71] van Wageningen, S. *et al.* Functional overlap and regulatory links shape genetic interactions between signaling pathways. *Cell* **143**, 991–1004 (2010).

[72] Linding, R. *et al.* Systematic discovery of in vivo phosphorylation networks. *Cell* **129**, 1415–1426 (2007).

[73] Sharifpoor, S. *et al.* A quantitative literature-curated gold standard for kinase-substrate pairs. *Genome Biology* **12**, R39 (2011).

[74] Boekhorst, J., Van Breukelen, B., Heck, A. & Snel, B. Comparative phosphoproteomics reveals evolutionary and functional conservation of phosphorylation across eukaryotes. *Genome Biology* **9**, R144 (2008).

[75] Landry, C., Levy, E. & Michnick, S. Weak functional constraints on phosphoproteomes. *Trends in Genetics* **25**, 193–197 (2009).

[76] Beltrao, P. *et al.* Evolution of phosphoregulation: comparison of phosphorylation patterns across yeast species. *PLOS Biology* **7**, e1000134 (2009).

[77] Tan, C. *et al.* Comparative analysis reveals conserved protein phosphorylation networks implicated in multiple diseases. *Science Signaling* **2**, ra39 (2009).

[78] Ba, A. & Moses, A. Evolution of characterized phosphorylation sites in budding yeast. *Molecular Biology and Evolution* **27**, 2027–2037 (2010).

[79] Holt, L. *et al.* Global analysis of Cdk1 substrate phosphorylation sites provides insights into evolution. *Science* **325**, 1682–1686 (2009).

[80] Freschi, L., Courcelles, M., Thibault, P., Michnick, S. & Landry, C. Phosphorylation network rewiring by gene duplication. *Molecular Systems Biology* **7** (2011).

[81] Brusniak, M. *et al.* Corra: Computational framework and tools for LC-MS discovery and targeted mass spectrometry-based proteomics. *BMC Bioinformatics* **9**, 542 (2008).

[82] Smyth, G. *et al.* Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology* **3**, 3 (2004).

[83] Bodenmiller, B. *et al.* Phosphoproteomic analysis reveals interconnected system-wide responses to perturbations of kinases and phosphatases in yeast. *Science Signaling* **3**, rs4 (2010).

[84] Kholodenko, B. N. Cell-signalling dynamics in time and space. *Nature Reviews Molecular Cell Biology* **7**, 165–176 (2006).

[85] Jordan, J. D., Landau, E. M. & Iyengar, R. Signaling networks: the origins of cellular multitasking. *Cell* **103**, 193–200 (2000).

[86] Hunter, T. Signaling–2000 and beyond. *Cell* **100**, 113–127 (2000).

[87] Kitano, H. Biological robustness. *Nature Reviews Genetics* **5**, 826–837 (2004).

[88] Félix, M. & Wagner, A. Robustness and evolution: concepts, insights and challenges from a developmental model system. *Heredity* **100**, 132–140 (2006).

[89] Schwartz, D. & Gygi, S. An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets. *Nature Biotechnology* **23**, 1391–1398 (2005).

[90] Mok, J. *et al.* Deciphering protein kinase specificity through large-scale analysis of yeast phosphorylation site motifs. *Science Signaling* **3**, ra12 (2010).

[91] Costanzo, M. *et al.* The genetic landscape of a cell. *Science* **327**, 425–431 (2010).

[92] Ashburner, M. *et al.* Gene Ontology: tool for the unification of biology. *Nature Genetics* **25**, 25–29 (2000).

[93] Unanchored dendrogram of *S.cerevisae* protein kinases (1997). URL `http://kinase.com/scerevisiae/`.

[94] Szklarczyk, D. *et al.* The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Research* **39**, D561 (2011).

[95] Stark, C. *et al.* The bioGRID interaction database: 2011 update. *Nucleic Acids Research* **39**, D698 (2011).

[96] Fiehn, O. Metabolomics - the link between genotypes and phenotypes. *Plant Molecular Biology* **48**, 155–171 (2002).

[97] Goodacre, R., Vaidyanathan, S., Dunn, W., Harrigan, G. & Kell, D. Metabolomics by numbers: acquiring and understanding global metabolite data. *Trends in Biotechnology* **22**, 245–252 (2004).

[98] Saito, K. & Matsuda, F. Metabolomics for Functional Genomics, Systems Biology, and Biotechnology. *Annual Review of Plant Biology* **61**, 463–489 (2010).

[99] Wishart, D. S. *et al.* HMDB: a knowledgebase for the human metabolome. *Nucleic Acids Research* **37**, D603–10 (2009).

[100] Dobson, P. D. *et al.* Further developments towards a genome-scale metabolic model of yeast. *BMC Systems Biology* **4**, 145 (2010).

[101] Fiehn, O. Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks. *Comparative and Functional Genomics* **2**, 155–168 (2001).

[102] Schmidt, C. Metabolomics: what's happening downstream of DNA. *Environmental Health Perspectives* **112**, A410 (2004).

[103] Kaddurah-Daouk, R., Kristal, B. S. & Weinshilboum, R. M. Metabolomics: a global biochemical approach to drug response and disease. *Annual Review of Pharmacology and Toxicology* **48**, 653–683 (2008).

[104] Saghatelian, A. & Cravatt, B. F. Global strategies to integrate the proteome and metabolome. *Current Opinion in Chemical Biology* **9**, 62–68 (2005).

[105] Beckmann, M., Parker, D., Enot, D., Duval, E. & Draper, J. High-throughput, nontargeted metabolite fingerprinting using nominal mass flow injection electrospray mass spectrometry. *Nature Protocols* **3**, 486–504 (2008).

[106] Fuhrer, T., Heer, D., Begemann, B. & Zamboni, N. High-throughput, accurate mass metabolome profiling of cellular extracts by flow injection-time-of-flight mass spectrometry. *Analytical Chemistry* (2011).

[107] Baker, M. Metabolomics: from small molecules to big ideas. *Nature Methods* **8**, 117–121 (2011).

[108] Mo, M. & Palsson, B. Understanding human metabolic physiology: a genome-to-systems approach. *Trends in Biotechnology* **27**, 37–44 (2009).

[109] Kell, D. Metabolomics and systems biology: making sense of the soup. *Current Opinion in Microbiology* **7**, 296–307 (2004).

[110] Weckwerth, W. & Fiehn, O. Can we discover novel pathways using metabolomic analysis? *Current Opinion in Biotechnology* **13**, 156–160 (2002).

[111] Bino, R. *et al.* Potential of metabolomics as a functional genomics tool. *Trends in Plant Science* **9**, 418–425 (2004).

[112] Henneges, C. *et al.* Prediction of breast cancer by profiling of urinary RNA metabolites using SVM-based feature selection. *BMC Cancer* **9**, 104 (2009).

[113] Li, X. *et al.* Comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry for metabonomics: Biomarker discovery for diabetes mellitus. *Analytica chimica acta* **633**, 257–262 (2009).

[114] Kaddurah-Daouk, R. *et al.* Metabolomic changes in autopsy-confirmed Alzheimer's disease. *Alzheimer's and Dementia* **7**, 309–317 (2011).

[115] Robertson, D. G. Metabonomics in toxicology: a review. *Toxicological Sciences* **85**, 809–822 (2005).

[116] Meyer, J. & Ginsburg, G. The path to personalized medicine. *Current Opinion in Chemical Biology* **6**, 434–438 (2002).

[117] van der Greef, J., Hankemeier, T. & McBurney, R. Metabolomics-based systems biology and personalized medicine: moving towards n = 1 clinical trials? *Pharmacogenomics* **7**, 1087–1094 (2006).

[118] Dettmer, K., Aronov, P. & Hammock, B. Mass spectrometry-based metabolomics. *Mass Spectrometry Reviews* **26**, 51 (2007).

[119] ter Kuile, B. & Westerhoff, H. Transcriptome meets metabolome: hierarchical and metabolic regulation of the glycolytic pathway. *FEBS letters* **500**, 169–171 (2001).

[120] Rossell, S., Weijden, C., Kruckeberg, A., Bakker, B. & Westerhoff, H. Hierarchical and metabolic regulation of glucose influx in starved Saccharomyces cerevisiae. *FEMS Yeast Research* **5**, 611–619 (2005).

[121] Ihmels, J., Levy, R., Barkai, N. *et al.* Principles of transcriptional control in the metabolic network of Saccharomyces cerevisiae. *Nature Biotechnology* **22**, 86–92 (2004).

[122] Patil, K. & Nielsen, J. Uncovering transcriptional regulation of metabolism by using metabolic network topology. *Proceedings of the National Academy of Sciences* **102**, 2685 (2005).

[123] Wei, H. *et al.* Transcriptional coordination of the metabolic network in Arabidopsis. *Plant Physiology* **142**, 762–774 (2006).

[124] Zhao, S. *et al.* Regulation of cellular metabolism by protein lysine acetylation. *Science* **327**, 1000–1004 (2010).

[125] Wang, Q. *et al.* Acetylation of metabolic enzymes coordinates carbon source utilization and metabolic flux. *Science* **327**, 1004–840 (2010).

[126] Gerosa, L. & Sauer, U. Regulation and control of metabolic fluxes in microbes. *Current Opinion in Biotechnology* 566–575 (2011).

[127] Norvell, A. & McMahon, S. Rise of the rival. *Science* **327**, 964 (2010).

[128] Sweetlove, L. & Fernie, A. Regulation of metabolic networks: understanding metabolic complexity in the systems biology era. *New Phytologist* **168**, 9–24 (2005).

[129] Schulz, J., Zampieri, M., Wanka, S., von Mering, C. & Sauer, U. Mapping the protein phosphorylation network that regulates metabolic activity in Saccharomyces cerevisiae. Manuscript in preparation.

[130] Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research* **28**, 27–30 (2000).

[131] KEGG compound C00043 *UDP-N-acetyl-D-glucosamine*. URL `http://www.genome.jp/dbget-bin/www_bget?cpd:C00043`.

[132] KEGG pathway sce00520 *Amino sugar and nucleotide sugar metabolism*. URL `http://www.genome.jp/dbget-bin/www_bget?pathway+sce00520`.

[133] Dittmar, G., Wilkinson, C., Jedrzejewski, P. & Finley, D. Role of a ubiquitin-like modification in polarized morphogenesis. *Science* **295**, 2442–2446 (2002).

[134] KEGG compound C00106 *Uracil*. URL `http://www.genome.jp/dbget-bin/www_bget?cpd:C00106`.

[135] KEGG pathway sce00410 *beta-Alanine metabolism*. URL `http://www.genome.jp/dbget-bin/www_bget?sce00410`.

[136] Liu, H. *et al.* The NOT proteins are part of the CCR4 transcriptional complex and affect gene expression both positively and negatively. *The EMBO journal* **17**, 1096–1106 (1998).

[137] KEGG compound C00149 *(S)-Malate*. URL `http://www.genome.jp/dbget-bin/www_bget?cpd:C00149`.

[138] KEGG pathway sce00680 *Methane metabolism*. URL `http://www.genome.jp/dbget-bin/www_bget?sce00680`.

[139] Kuhn, M. *et al.* STITCH 2: an interaction network database for small molecules and proteins. *Nucleic Acids Research* **38**, D552 (2010).

[140] KEGG compound C00921 *7,8-Dihydropteroate*. URL `http://www.genome.jp/dbget-bin/www_bget?cpd:C00921`.

[141] KEGG pathway sce00790 *Folate biosynthesis*. URL `http://www.genome.jp/dbget-bin/www_bget?pathway+sce00790`.

[142] KEGG compound C00024 *Acetyl-CoA*. URL `http://www.genome.jp/dbget-bin/www_bget?cpd:C00024`.

[143] KEGG compound C00332 *Acetoacetyl-CoA*. URL http://www.genome.jp/dbget-bin/www_bget?cpd:C00332.

[144] KEGG compound C00010 *Coenzyme A*. URL http://www.genome.jp/dbget-bin/www_bget?cpd:C00010.

[145] KEGG pathway sce00061 *Fatty acid biosynthesis*. URL http://www.genome.jp/dbget-bin/www_bget?pathway+sce00061.

[146] KEGG pathway sce00071 *Fatty acid metabolism*. URL http://www.genome.jp/dbget-bin/www_bget?pathway+sce00071.

[147] Hasslacher, M., Ivessa, A., Paltauf, F. & Kohlwein, S. Acetyl-coa carboxylase from yeast is an essential enzyme and is regulated by factors that control phospholipid metabolism. *Journal of Biological Chemistry* **268**, 10946 (1993).

[148] Usaite, R. *et al.* Reconstruction of the yeast Snf1 kinase regulatory network reveals its role as a global energy regulator. *Molecular Systems Biology* **5** (2009).

[149] Woods, A. *et al.* Yeast SNF1 is functionally related to mammalian AMP-activated protein kinase and regulates acetyl-CoA carboxylase in vivo. *Journal of Biological Chemistry* **269**, 19509–19515 (1994).

[150] Hovik, R., Brodal, B., Bartlett, K. & Osmundsen, H. Metabolism of acetyl-CoA by isolated peroxisomal fractions: formation of acetate and acetoacetyl-CoA. *Journal of Lipid Research* **32**, 993–999 (1991).

[151] Fatland, B. *et al.* Molecular characterization of a heteromeric ATP-citrate lyase that generates cytosolic acetyl-coenzyme A in Arabidopsis. *Plant Physiology* **130**, 740–756 (2002).